Table 9: GBS genes shared with pneumoccocus

ORFxxxxx Annotation
ORF01580 polysaccharide biosynthesis protein, putative
ORF01612 conserved hypothetical protein
ORF01613 glycosyl transferase, group 1 family protein
ORF01617 conserved hypothetical protein
ORF01618 conserved hypothetical protein
ORF01621 glycosyl transferase, putative
ORF01622 glycosyl transferase, group 2 family protein
ORF01623 glycosyl transferase, family 8, degenerate
ORF01624 IS1381, transposase OrfB
ORF01625 IS1381, transposase OrfA
ORF01626 glycosyl transferase family 8
ORF01627 glycosyl transferase, family 8
ORF01628 conserved hypothetical protein
ORF01630 cell wall surface anchor family protein
ORF01635 protease, putative
ORF01643 aminopeptidase PepS (pepS)
ORF01702 peptidase, M20/M25/M40 family
ORF01731 IS1381, transposase OrfA
ORF01732 IS1381, transposase OrfB
ORF01740 tellurite resistance protein TehB (tehB)
ORF01747 methylated-DNA--protein-cysteine S-methyltransferase (ogt)
ORF01749 acetyltransferase, GNAT family
ORF01763 AcuB family protein
ORF01764 branched-chain amino acid ABC transporter, ATP-binding protein (livF)
ORF01765 branched-chain amino acid ABC transporter, ATP-binding protein (livG)
ORF01766 branched-chain amino acid ABC transporter, permease protein
ORF01767 branched-chain amino acid ABC transporter, permease protein (livH)
ORF01769 branched-chain amino acid ABC transporter, amino acid-binding protein
ORF01775 aminotransferase, class I
ORF01779 potassium uptake protein, Trk family
ORF01780 cation uptake protein, Trk family
ORF01824 cobalt transport family protein
ORF01826 conserved hypothetical protein
ORF01832 peptidase, M20/M25/M40 family
ORF01845 conserved hypothetical protein
ORF01848 transcriptional regulator, MerR family
ORF01853 isochorismatase family protein
ORF01859 membrane protein
ORF01875 oxidoreductase, aldo/keto reductase family
ORF01880 phospho-2-dehydro-3-deoxyheptonate aldolase
ORF01981 rRNA (guanine-N1-)-methyltransferase, putative
ORF02083 prophage LambdaSa2, DNA replication protein DnaC, putative
ORF02101 Na+/H+ exchanger family protein
ORF02107 membrane protein, putative
ORF02139 UDP-glucose 4-epimerase (galE)
ORF02143 lacX protein
ORF02162 conserved hypothetical protein
ORF02186 hemolysin precursor, putative
ORF02192 transcriptional regulator, MerR family
ORF02195 MutT/nudix family protein
ORF02228 IS1381, transposase OrfB
ORF02229 IS1381, transposase OrfA
ORF02233 conserved hypothetical protein
ORF02234 conserved hypothetical protein
ORF02276 5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase (metE)

Table 9: GBS genes shared with pneumoccocus

ORFxxxxx Annotation
ORF02278 branched-chain amino acid transport protein AzlC, putative
ORF02288 glycosyl transferase, family 8
ORF02289 glycosyl transferase, family 8
ORF02341 ribosomal protein L32 (rpmF)
ORF02343 conserved hypothetical protein
ORF02358 sensor histidine kinase
ORF02369 conserved hypothetical protein
ORF02384 LysM domain protein
ORF02428 hypoxanthine-guanine phosphoribosyltransferase (hpt)
ORF03011 ribosomal protein L33
ORF03014 ribosomal protein L33

Table 10: GBS genes shared with GAS

ORFxxxxx	Annotation
ORF00064	ribosomal protein S14, putative
ORF00095	D-alanyl-D-alanine carboxypeptidase family protein
ORF00096	N-acetylmuramoyl-L-alanine amidase, family 4 protein
ORF00110	conserved hypothetical protein
ORF00112	DNA repair protein RadA (radA)
ORF00124	permease, putative
ORF00148	glycosyl transferase, group 4 family protein
ORF00154	penicillin-binding protein 4, putative
ORF00157	oligopeptide ABC transporter, permease protein
ORF00206	oligopeptide ABC transporter, oligopeptide-binding protein
ORF00207	oligopeptide ABC transporter, permease protein
ORF00208	oligopeptide ABC transporter, permease protein
ORF00209	peptide ABC transporter, ATP-binding protein
ORF00210	peptide ABC transporter, ATP-binding protein
ORF00216	IS1548, transposase
ORF00226	conserved hypothetical protein
ORF00232	conserved hypothetical protein
ORF00239	site-specific recombinase, phage integrase family
ORF00250	conserved hypothetical protein
ORF00251	conserved hypothetical protein
ORF00289	ABC transporter, ATP-binding protein
ORF00305	NADH oxidase, putative
ORF00317	cell division protein FtsL, putative
ORF00333	conserved hypothetical protein
ORF00383	hydrolase, haloacid dehalogenase-like family
ORF00430	expressed putative lipoprotein
ORF00431	transcriptional repressor CopY
ORF00434	membrane protein, putative
ORF00438	transcriptional regulator, Fur family
ORF00442	membrane protein, putative
ORF00445	bioY family protein
ORF00446	AtsA/ElaC family protein
ORF00468	expressed putative protease
ORF00469	glycosyl transferase, group 2 family protein
ORF00471	nrdI protein (nrdI)
ORF00473	expressed protein of unknown function
ORF00474	conserved hypothetical protein
ORF00507	conserved hypothetical protein
ORF00525	bioY family protein
ORF00528	thiolase
ORF00531	AMP-binding enzyme domain protein
ORF00548	YGGT family protein
ORF00565	exodeoxyribonuclease VII, small subunit (xseB)
ORF00568	arginine repressor ArgR, putative
ORF00572	expressed putative lipase/acylhydrolase
ORF00573	conserved hypothetical protein
ORF00586	iron-sulfur cluster-binding protein, putative
ORF00592	oxidoreductase, short chain dehydrogenase/reductase family
ORF00604	dipeptidase
ORF00611	voltage-gated chloride channel family protein
ORF00619	prophage LambdaSa1, repressor protein, putative
ORF00622	conserved hypothetical protein
ORF00627	prophage LambdaSa1, antirepressor, putative
ORF00634	conserved hypothetical protein
ORF00648	conserved hypothetical protein

Table 10: GBS genes shared with GAS

ORFxxxxx	Annotation
ORF00654	conserved hypothetical protein
ORF00655	conserved hypothetical protein
ORF00656	conserved hypothetical protein
ORF00658	conserved hypothetical protein
ORF00659	conserved hypothetical protein
ORF00660	prophage LambdaSa1, structural protein, putative
ORF00662	conserved hypothetical protein
ORF00663	conserved hypothetical protein
ORF00664	conserved hypothetical protein
ORF00665	conserved hypothetical protein
ORF00666	prophage LambdaSa1, structural protein
ORF00668	conserved hypothetical protein
ORF00669	prophage LambdaSa1, pblA protein, internal deletion
ORF00677	prophage LambdaSa1, lysin, putative
ORF00679	conserved hypothetical protein
ORF00695	transposase OrfB, IS3 family, truncation
ORF00697	conserved hypothetical protein
ORF00707	conserved domain protein
ORF00713	acid phosphatase precursor, class B
ORF00720	transposase OrfB, IS3 family FRAMESHIFT
ORF00721	transposase OrfA, IS3 family
ORF00751	cylA protein (cylA)
ORF00755	cylI protein (cylI)
ORF00760	serine protease, subtilase family, putative POINT MUTATION
ORF00781	transcriptional regulator, LysR family
ORF00783	regulatory protein, putative
ORF00785	IS1548, transposase
ORF00786	regulatory protein, putative, truncation
ORF00787	D-lactate dehydrogenase (ldhA)
ORF00801	glycosyl transferase, group 1 family protein
ORF00805	conserved hypothetical protein
ORF00826	phage shock protein C, putative
ORF00833	conserved hypothetical protein
ORF00845	hydrolase, haloacid dehalogenase-like family
ORF00852	conserved hypothetical protein
ORF00853	expressed putative lipoprotein
ORF00857	IS1548, transposase
ORF00890	conserved hypothetical protein
ORF00902	conserved hypothetical protein
ORF00926	membrane protein, putative
ORF00927	membrane protein, putative
ORF00987	conserved hypothetical protein
ORF01009	expressed protein of unknown function
ORF01010	lipoyl-binding domain protein
ORF01011	oxidoreductase, putative
ORF01012	conserved hypothetical protein
ORF01024	expressed putative lipoprotein
ORF01061	signal peptidase I, putative
ORF01064	IS1548, transposase
ORF01084	glyoxylase family protein
ORF01104	SatD
ORF01126	conserved hypothetical protein
ORF01191	conserved hypothetical protein
ORF01192	conserved hypothetical protein
ORF01193	glycine cleavage system H protein, putative

Table 10: GBS genes shared with GAS

ORFxxxxx Annotation
ORF01194 bacterial luciferase family protein
ORF01195 oxidoreductase, FMN-binding
ORF01197 lipoate-protein ligase A family protein
ORF01202 IS861, transposase OrfA
ORF01223 drug resistance transporter, EmrB/QacA family, putative
ORF01224 conserved hypothetical protein
ORF01225 potassium uptake protein, putative
ORF01237 membrane protein, putative
ORF01249 dihydroneopterin aldolase (folB)
ORF01256 polysaccharide deacetylase family protein
ORF01273 transcriptional regulator, GntR family/potassioum uptake protein, TrkA family
ORF01280 conserved hypothetical protein
ORF01281 conserved hypothetical protein
ORF01289 lipoprotein, putative
ORF01291 conserved hypothetical protein
ORF01298 conserved hypothetical protein
ORF01318 conserved hypothetical protein
ORF01320 voltage-gated chloride channel family protein, putative
ORF01322 arsenate reductase (arsC)
ORF01340 dTDP-glucose 4,6-dehydratase (rfbB)
ORF01341 dTDP-4-dehydrorhamnose 3,5-epimerase
ORF01342 glucose-1-phosphate thymidylyltransferase (rfbA)
ORF01356 hypothetical protein
ORF01368 conserved hypothetical protein
ORF01374 ISSdy1, transposase OrfB
ORF01388 transposase OrfA, IS3 family
ORF01389 transposase OrfB, IS3 family, truncation
ORF01391 ISSdy1, transposase OrfB FRAMESHIFT
ORF01396 transcriptional regulator, Cro/CI family
ORF01419 repressor protein, putative
ORF01461 amino acid permease
ORF01469 conserved hypothetical protein
ORF01483 sensor histidine kinase
ORF01485 GTP pyrophosphokinase family protein
ORF01490 5'-nucleotidase family protein
ORF01509 2-dehydropantoate 2-reductase, putative
ORF01510 regulatory protein, putative
ORF01522 carbamoyl-phosphate synthase, large subunit, putative
ORF01542 sulfatase
ORF01549 conserved hypothetical protein
ORF01554 iron compound ABC transporter, substrate-binding protein
ORF01557 conserved hypothetical protein
ORF01563 conserved hypothetical protein TIGR01212
ORF01583 glycosyltransferase, group 2 family protein
ORF01584 glycosyltransferase, group 2 family protein
ORF01585 glycosyltransferase, putative
ORF01586 dTDP-4-dehydrorhamnose reductase (rfbD)
ORF01593 conserved hypothetical protein
ORF01599 conserved hypothetical protein
ORF01600 glycerol-3-phosphate transporter, putative
ORF01639 conserved hypothetical protein
ORF01650 nitroreductase family protein
ORF01653 amino acid permease
ORF01665 transcriptional regulator, MutR family
ORF01683 MutT/nudix family protein

Table 10: GBS genes shared with GAS

ORFxxxxx	**Annotation**
ORF01686	67 kDa Myosin-crossreactive streptococcal antigen
ORF01688	peptide methionine sulfoxide reductase (msrA)
ORF01694	peptide ABC transporter, permease protein
ORF01704	conserved hypothetical protein
ORF01705	IS861, transposase OrfA
ORF01741	membrane protein, putative
ORF01770	conserved hypothetical protein
ORF01772	IS1548, transposase
ORF01790	conserved hypothetical protein
ORF01794	conserved hypothetical protein
ORF01800	amino acid ABC transporter, substrate-binding protein
ORF01810	IS1548, transposase
ORF01827	sodium:dicarboxylate symporter family protein
ORF01877	immunogenic secreted protein, putative
ORF01913	transcriptional regulator, Cro/CI family
ORF01928	membrane protein, putative
ORF01931	transporter, putative
ORF01932	transcriptional regulator, Crp/Fnr family
ORF01947	transcriptional regulator, merR family
ORF01970	acid phosphatase
ORF02002	amino acid ABC transporter, permease protein
ORF02028	perfringolysin O regulator protein (pfoR)
ORF02029	conserved hypothetical protein
ORF02031	expressed protein of unknown function
ORF02032	expressed protein of unknown function
ORF02035	deoxynucleoside kinase family protein
ORF02042	alkyl hydroperoxide reductase, subunit C (ahpC)
ORF02126	transcriptional regulator, MarR family
ORF02128	N-acetylmuramoyl-L-alanine amidase, family 4 protein
ORF02135	malate oxidoreductase
ORF02136	citrate carrier protein, CCS family
ORF02137	sensor histidine kinase family protein
ORF02138	response regulator
ORF02166	conserved hypothetical protein
ORF02169	PTS system, IIB component
ORF02170	PTS system, IIA component, putative
ORF02202	ABC transporter, ATP-binding protein
ORF02262	ABC transporter, ATP-binding protein
ORF02270	cAMP factor (cfb)
ORF02280	serine protease, subtilase family, putative
ORF02286	major facilitator family protein
ORF02292	preprotein translocase, SecE subunit, putative
ORF02295	Lyme disease proteins of unknown function, putative
ORF02298	Na+ dependent nucleoside transporter
ORF02301	transcriptional regulator, GntR family
ORF02313	virulence factor MviM, putative
ORF02316	membrane protein, putative
ORF02319	conserved hypothetical protein TIGR00250
ORF02328	transporter, putative
ORF02331	cold shock protein, CSD family
ORF02332	DNA mismatch repair protein HexA (hexA)
ORF02335	conserved hypothetical protein
ORF02372	conserved hypothetical protein
ORF02383	expressed putative lipoprotein
ORF02393	transporter, putative

Table 10: GBS genes shared with GAS

ORFxxxxx	Annotation
ORF02398	transcriptional regulator, Crp/Fnr family
ORF02399	conserved hypothetical protein
ORF02401	acetyltransferase, GNAT family
ORF02403	arginine/ornithine antiporter (arcD)
ORF03002	conserved hypothetical protein, truncation

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF00008 protease, putative
ORF00010 acyl carrier protein (acpP)
ORF00016 acetyltransferase, GNAT family
ORF00018 peptidase, M23/M37 family, putative secreted protein
ORF00035 membrane protein, putative
ORF00087 lipoprotein, putative
ORF00088 hypothetical protein
ORF00089 hypothetical protein
ORF00091 conserved hypothetical protein
ORF00117 ribose ABC transporter, periplasmic D-ribose-binding protein (rbsB)
ORF00118 ribose ABC transporter, permease protein (rbsC)
ORF00120 ribose ABC transporter protein RbsD (rbsD)
ORF00121 ribokinase (rbsK)
ORF00123 hypothetical protein
ORF00130 argininosuccinate lyase (argH)
ORF00137 conserved hypothetical protein
ORF00138 hypothetical protein
ORF00166 4-diphosphocytidyl-2C-methyl-D-erythritol kinase (ispE)
ORF00182 conserved domain protein
ORF00186 transcriptional regulator, Cro/CI family
ORF00187 hypothetical protein
ORF00188 hypothetical protein
ORF00192 hypothetical protein
ORF00193 conserved hypothetical protein
ORF00196 conserved hypothetical protein
ORF00199 hydrolase, haloacid dehalogenase-like family
ORF00200 sensor histidine kinase, putative
ORF00201 response regulator
ORF00203 conserved hypothetical protein
ORF00204 membrane protein, putative
ORF00205 hypothetical protein
ORF00228 lipoprotein, putative
ORF00234 hypothetical protein
ORF00235 hypothetical protein
ORF00238 hypothetical protein
ORF00240 transcriptional regulator, Cro/CI family
ORF00241 hypothetical protein
ORF00242 conserved hypothetical protein
ORF00243 hypothetical protein
ORF00244 conserved domain protein
ORF00245 conserved hypothetical protein, fusion
ORF00246 replication initiation protein, putative
ORF00247 hypothetical protein
ORF00248 recombination protein
ORF00249 hypothetical protein
ORF00252 conserved hypothetical protein
ORF00253 hypothetical protein
ORF00254 hypothetical protein
ORF00255 hypothetical protein
ORF00256 hypothetical protein
ORF00257 hypothetical protein
ORF00258 hypothetical protein
ORF00259 hypothetical protein
ORF00260 hypothetical protein
ORF00272 expressed putative lipoprotein

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF00273 hypothetical protein
ORF00274 hypothetical protein
ORF00275 hypothetical protein
ORF00276 hypothetical protein
ORF00278 membrane protein, putative
ORF00279 transcriptional regulator, Cro/CI family
ORF00280 acetyltransferase, GNAT family
ORF00281 acetyltransferase, GNAT family
ORF00283 conserved hypothetical protein
ORF00284 RNA polymerase sigma factor, ECF subfamily
ORF00285 lipoprotein, putative
ORF00287 transcriptional regulator, TetR family
ORF00288 ABC transporter efflux protein, DrrB family, putative
ORF00292 hypothetical protein
ORF00294 expressed protein of unknown function
ORF00298 acyl carrier protein phosphodiesterase, putative
ORF00308 conserved hypothetical protein
ORF00324 conserved hypothetical protein
ORF00332 hypothetical protein
ORF00340 hypothetical protein
ORF00347 conserved hypothetical protein
ORF00384 hypothetical protein
ORF00402 membrane protein, putative
ORF00408 hypothetical protein
ORF00409 membrane protein, putative
ORF00414 conserved hypothetical protein
ORF00416 hypothetical protein
ORF00417 hypothetical protein
ORF00433 copper-transporter protein CopZ
ORF00448 hypothetical protein
ORF00466 conserved hypothetical protein
ORF00467 acetyltransferase, GNAT family
ORF00475 conserved domain protein
ORF00476 hypothetical protein
ORF00478 carboxymuconolactone decarboxylase family protein
ORF00479 conserved hypothetical protein
ORF00486 transcriptional regulator, AraC family
ORF00487 surface protein Rib
ORF00488 transposase, IS256 family, truncation
ORF00489 DNA-damage-inducible protein J, putative
ORF00490 hypothetical protein
ORF00491 lipoprotein, putative
ORF00493 bacteriophage L54a, integrase, truncation
ORF00497 conserved domain protein
ORF00503 oxidoreductase, Gfo/Idh/MocA family
ORF00506 transposase, IS256 family
ORF00510 bacteriocin transport accessory protein,putative
ORF00512 hypothetical protein
ORF00526 biotin synthetase (bioB)
ORF00527 hypothetical protein
ORF00533 type IV prepilin peptidase-related protein
ORF00538 conserved hypothetical protein
ORF00556 hypothetical protein
ORF00563 expressed protein of unknown function
ORF00575 hypothetical protein

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF00584 conserved hypothetical protein
ORF00585 fructose-1,6-bisphosphatase, putative
ORF00590 carboxymethylenebutenolidase-related protein
ORF00597 conserved hypothetical protein
ORF00598 inosine-uridine preferring nucleoside hydrolase
ORF00599 hypothetical protein
ORF00600 OsmC/Ohr family protein
ORF00608 adenosine deaminase, putative
ORF00610 chorismate mutase, putative
ORF00615 prophage LambdaSa1, site-specific recombinase, phage integrase family
ORF00617 conserved domain protein
ORF00618 hypothetical protein
ORF00620 hypothetical protein
ORF00621 conserved hypothetical protein
ORF00623 hypothetical protein
ORF00624 hypothetical protein
ORF00626 prophage LambdaSa1, transcriptional regulator, Cro/CI family
ORF00628 hypothetical protein
ORF00630 hypothetical protein
ORF00632 hypothetical protein
ORF00633 conserved hypothetical protein
ORF00635 hypothetical protein
ORF00636 hypothetical protein
ORF00637 hypothetical protein
ORF00638 conserved hypothetical protein
ORF00639 conserved domain protein
ORF00641 prophage LambdaSa1, reverse transcriptase/maturase family protein
ORF00642 conserved hypothetical protein
ORF00643 conserved hypothetical protein
ORF00644 hypothetical protein
ORF00645 hypothetical protein
ORF00646 conserved hypothetical protein
ORF00647 hypothetical protein
ORF00649 hypothetical protein
ORF00650 hypothetical protein
ORF00652 conserved hypothetical protein
ORF00653 conserved hypothetical protein
ORF00657 conserved hypothetical protein, truncation
ORF00661 conserved hypothetical protein
ORF00667 conserved hypothetical protein
ORF00670 prophage LambdaSa1, minor structural protein, putative
ORF00671 prophage LambdaSa1, N-acetylmuramoyl-L-alanine amidase, family 4
ORF00672 prophage LambdaSa1, minor structural protein, putative
ORF00673 hypothetical protein
ORF00674 hypothetical protein
ORF00675 conserved hypothetical protein
ORF00676 conserved hypothetical protein
ORF00678 conserved hypothetical protein
ORF00681 conserved hypothetical protein
ORF00682 hypothetical protein
ORF00683 prophage LambdaSa1, site-specific recombinase, phage integrase family FRAMESHIFT
ORF00685 conserved hypothetical protein
ORF00689 conserved hypothetical protein, FRAMESHIFT
ORF00698 hypothetical protein
ORF00703 phosphoserine phosphatase SerB (serB)

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF00704 MutT/nudix family protein
ORF00712 hypothetical protein
ORF00718 cell wall surface protein, interruption-N
ORF00723 hypothetical protein
ORF00726 transcriptional regulator, AraC family
ORF00727 expressed cell wall surface anchor family protein
ORF00728 expressed cell wall surface anchor family protein
ORF00735 expressed protein of unknown function
ORF00737 conserved hypothetical protein, degenerate
ORF00738 hypothetical protein
ORF00740 hypothetical protein
ORF00741 hypothetical protein
ORF00742 lipoprotein, putative
ORF00747 cylD protein (cylD)
ORF00749 acyl carrier protein AcpC
ORF00750 cylZ protein FRAMESHIFT
ORF00752 cylB protein (cylB)
ORF00753 cylE protein (cylE)
ORF00754 cylF protein (cylF)
ORF00756 cylJ protein (cylJ)
ORF00757 cylK protein (cylK)
ORF00758 hypothetical protein
ORF00759 putative secreted protein
ORF00761 hypothetical protein
ORF00766 expressed putative secreted protein
ORF00767 hypothetical protein
ORF00768 conserved domain protein
ORF00769 permease, putative
ORF00775 conserved hypothetical protein
ORF00777 DedA family protein, putative
ORF00779 membrane protein, putative
ORF00788 sodium:galactoside symporter family protein, putative
ORF00791 transcriptional regulator, GntR family
ORF00793 Glucuronate isomerase (uxaC)
ORF00794 mannonate dehydratase (uxuA)
ORF00795 D-mannonate oxidoreductase
ORF00796 hydrolase, haloacid dehalogenase-like family
ORF00797 glycosyl hydrolase, family 3
ORF00806 conserved hypothetical protein
ORF00822 ABC transporter, ATP-binding protein
ORF00827 hypothetical protein
ORF00834 conserved hypothetical protein
ORF00838 membrane protein, putative
ORF00839 Mn2+/Fe2+ transporter, NRAMP family
ORF00848 conserved domain protein
ORF00872 cell wall surface anchor family protein
ORF00874 conserved hypothetical protein
ORF00878 ABC transporter, permease protein
ORF00879 YaeC family protein, putative
ORF00888 hydrolase, haloacid dehalogenase-like family
ORF00891 conserved domain protein
ORF00898 conserved hypothetical protein
ORF00900 permease, GntP family
ORF00903 transcriptional regulator, MarR family
ORF00907 glutathione S-transferase family protein

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx	Annotation
ORF00909	hypothetical protein
ORF00921	membrane protein, putative
ORF00922	glycosyl transferase, family 8
ORF00923	hypothetical protein
ORF00924	conserved hypothetical protein
ORF00939	conserved hypothetical protein
ORF00942	expressed putative secreted protein
ORF00943	hypothetical protein
ORF00944	hypothetical protein
ORF00946	conserved hypothetical protein
ORF00950	hypothetical protein
ORF00951	transcriptional regulator, TenA family
ORF00972	ATP synthase F0, C subunit (atpE)
ORF00980	conserved hypothetical protein
ORF00982	conserved hypothetical protein
ORF01003	conserved hypothetical protein
ORF01004	conserved hypothetical protein
ORF01013	hypothetical protein
ORF01014	hypothetical protein
ORF01015	hypothetical protein
ORF01016	hypothetical protein
ORF01018	hypothetical protein
ORF01019	hypothetical protein
ORF01021	hypothetical protein
ORF01025	HD domain protein
ORF01026	acetyltransferase, GNAT family
ORF01032	chloramphenicol acetyltransferase (cat)
ORF01034	Tn916, transposase
ORF01035	Tn916, excisionase
ORF01037	Tn916, hypothetical protein
ORF01038	Tn916, hypothetical protein
ORF01039	Tn916, transcriptional regulator, putative
ORF01041	Tn916, hypothetical protein
ORF01042	Tn916, NLP/P60 family protein
ORF01044	membrane protein, putative FRAMESHIFT
ORF01048	Tn916, hypothetical protein
ORF01049	Tn916, hypothetical protein
ORF01050	Tn916, hypothetical protein
ORF01051	Tn916, transcriptional regulator, putative
ORF01052	Tn916, FtsK/SpoIIIE family protein
ORF01053	Tn916, hypothetical protein
ORF01054	Tn916, hypothetical protein
ORF01062	hypothetical protein
ORF01086	Na+/H+ exchanger family protein
ORF01092	acetyltransferase, GNAT family
ORF01096	nisin-resistance protein, putative
ORF01103	conserved hypothetical protein
ORF01124	acetyltransferase, GNAT family
ORF01133	iron-compound ABC transporter, iron-compound-binding protein
ORF01140	conserved hypothetical protein
ORF01142	carbon starvation protein CstA, putative
ORF01143	response regulator
ORF01144	sensor histidine kinase, putative
ORF01145	lipoprotein, putative
ORF01146	conserved hypothetical protein, FRAMESHIFT

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF01148 lipoprotein, putative
ORF01149 hypothetical protein
ORF01150 hypothetical protein
ORF01151 hypothetical protein
ORF01152 lipoprotein, putative
ORF01153 hypothetical protein
ORF01157 conserved hypothetical protein
ORF01158 hypothetical protein
ORF01159 hypothetical protein
ORF01160 expressed protein of unknown function FRAMESHIFT
ORF01161 expressed conserved domain protein
ORF01162 conserved hypothetical protein
ORF01164 FtsK/SpoIIIE family protein FRAMESHIFT
ORF01166 hypothetical protein
ORF01167 conserved hypothetical protein
ORF01168 conserved hypothetical protein
ORF01169 hypothetical protein
ORF01172 phage infection protein, putative
ORF01173 conserved hypothetical protein
ORF01174 conserved domain protein
ORF01175 hypothetical protein
ORF01182 membrane protein, putative
ORF01186 cell wall surface anchor family protein, putative
ORF01187 hypothetical protein
ORF01204 hypothetical protein
ORF01215 hypothetical protein
ORF01241 transcriptional regulator, AraC family, putative
ORF01253 rarD protein (rarD)
ORF01257 transporter, BCCT family protein
ORF01258 hypothetical protein
ORF01261 expressed protein of unknown function
ORF01262 conserved hypothetical protein, FRAMESHIFT
ORF01263 hypothetical protein
ORF01265 hypothetical protein
ORF01266 hypothetical protein
ORF01269 conserved hypothetical protein
ORF01272 conserved hypothetical protein
ORF01277 conserved hypothetical protein
ORF01287 conserved hypothetical protein
ORF01288 membrane protein, putative
ORF01299 CMP-N-acetylneuraminic acid synthetase NeuA (neuA)
ORF01300 neuD protein (neuD)
ORF01301 UDP-N-acetylglucosamine-2-epimerase NeuC (neuC)
ORF01302 N-acetyl neuramic acid synthetase NeuB (neuB)
ORF01303 polysaccharide biosynthesis protein CpsL (cpsL)
ORF01304 polysaccharide biosynthesis protein CpsK(V) (cpsK)
ORF01307 glycosyltransferase CpsN(V) (cpsN)
ORF01308 polysaccharide biosynthesis protein CpsM(V) (cpsM)
ORF01309 polysaccharide biosynthesis protein cpsH(V) (cpsH)
ORF01310 glycosyltransferase CpsG(V) (cpsG)
ORF01311 polysaccharide biosynthesis protein CpsF (cpsF)
ORF01312 glycosyltransferase CpsE (cpsE)
ORF01348 conserved domain protein
ORF01349 hypothetical protein
ORF01370 conserved hypothetical protein

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF01371 conserved hypothetical protein
ORF01372 expressed protein of unknown function
ORF01373 ISSdy1, transposase OrfA
ORF01375 conserved hypothetical protein
ORF01379 transposase OrfB, IS3 family, truncation
ORF01382 GBSi1, group II intron, maturase
ORF01384 hypothetical protein
ORF01385 hypothetical protein
ORF01386 conserved hypothetical protein
ORF01387 conserved hypothetical protein, truncation
ORF01390 ISSdy1, transposase OrfA FRAMESHIFT
ORF01392 hypothetical protein
ORF01393 hypothetical protein
ORF01394 site-specific recombinase, phage integrase family
ORF01395 conserved hypothetical protein
ORF01401 transposase, ISL3 family
ORF01404 mercuric resistance operon regulatory protein MerR (merR)
ORF01408 cadmium efflux system accessory protein (CadC)
ORF01409 conserved hypothetical protein
ORF01410 hypothetical protein
ORF01417 hypothetical protein
ORF01418 hypothetical protein
ORF01420 hypothetical protein
ORF01421 ImpB/MucB/SamB family protein
ORF01423 conserved hypothetical protein
ORF01424 conserved hypothetical protein
ORF01425 conserved hypothetical protein
ORF01426 conserved hypothetical protein
ORF01427 hypothetical protein
ORF01428 conserved hypothetical protein
ORF01430 hypothetical protein
ORF01431 hypothetical protein
ORF01432 conserved domain protein
ORF01433 SNF2 family protein
ORF01434 hypothetical protein
ORF01435 calcium-binding protein, putative
ORF01436 agglutinin receptor (ssp-5)
ORF01437 abortive infection protein AbiGI (abiGI)
ORF01438 abortive infection protein AbiGII (abiGII)
ORF01439 conserved hypothetical protein
ORF01440 expressed protein of unknown function
ORF01441 conserved hypothetical protein, degenerate
ORF01442 membrane protein, putative
ORF01443 hypothetical protein
ORF01444 Tn5252, Orf 21 protein, internal deletion
ORF01445 hypothetical protein
ORF01450 conserved hypothetical protein
ORF01452 hypothetical protein
ORF01454 conserved hypothetical protein
ORF01459 hypothetical protein
ORF01460 homocysteine S-methyltransferase MmuM, putative
ORF01463 hypothetical protein
ORF01464 hypothetical protein
ORF01465 hypothetical protein
ORF01466 transcriptional regulator, TetR family

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF01477 glutathione S-transferase family protein, putative
ORF01478 conserved domain protein
ORF01486 hypothetical protein
ORF01488 R5 protein
ORF01489 transcriptional regulator, MarR family, putative
ORF01494 membrane protein, putative
ORF01497 acetyltransferase, GNAT family
ORF01502 hypothetical protein
ORF01503 conserved hypothetical protein
ORF01508 surface antigen-related protein
ORF01535 conserved hypothetical protein
ORF01547 conserved hypothetical protein
ORF01566 expressed cell wall surface anchor family protein
ORF01572 glycosyltransferase, group 1 family protein
ORF01573 glycosyltransferase, group 2 family protein
ORF01575 membrane protein, putative
ORF01576 glycosyltransferase, group 2 family protein
ORF01577 glycosyltransferase, group 2 family protein
ORF01578 nucleotide sugar dehydratase, putative
ORF01581 lipoprotein, putative
ORF01582 conserved hypothetical protein
ORF01596 ammonium transporter family protein
ORF01597 conserved hypothetical protein
ORF01601 hypothetical protein
ORF01608 proton/peptide symporter family protein
ORF01611 hypothetical protein
ORF01615 conserved domain protein
ORF01638 conserved hypothetical protein
ORF01641 conserved hypothetical protein
ORF01645 cell wall surface anchor family protein
ORF01660 membrane protein, putative
ORF01661 ABC transporter, ATP binding protein
ORF01666 hypothetical protein
ORF01667 hypothetical protein
ORF01670 hypothetical protein
ORF01672 protease, putative, POINT MUTATION
ORF01673 hypothetical protein
ORF01674 hypothetical protein
ORF01675 hypothetical protein
ORF01680 tetracenomycin polyketide synthesis O-methyltransferase TcmP, putative
ORF01681 hypothetical protein
ORF01682 hypothetical protein
ORF01684 hypothetical protein
ORF01692 peptide ABC transporter, ATP-binding protein
ORF01695 peptide ABC transporter, permease protein
ORF01696 peptide ABC transporter, peptide-binding protein
ORF01699 transposase, IS30 family, putative
ORF01700 transporter, major facilitator family
ORF01703 transcriptional regulator, LysR family
ORF01715 conserved hypothetical protein
ORF01719 hypothetical protein
ORF01720 conserved hypothetical protein
ORF01721 glyoxalase family protein
ORF01727 conserved hypothetical protein
ORF01729 acetyltransferase, GNAT family

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF01730 glycosyl transferase, group 2 family protein
ORF01733 hypothetical protein
ORF01734 conserved hypothetical protein
ORF01735 hypothetical protein
ORF01736 hypothetical protein
ORF01737 hypothetical protein
ORF01742 hypothetical protein
ORF01743 PTS system component, putative
ORF01744 conserved hypothetical protein
ORF01748 D-isomer specific 2-hydroxyacid dehydrogenase family protein
ORF01753 conserved hypothetical protein
ORF01754 hypothetical protein
ORF01761 transposase, IS30 family, putative, truncation
ORF01778 amino acid permease, putative
ORF01807 hypothetical protein
ORF01836 hypothetical protein
ORF01838 hypothetical protein
ORF01839 dihydroxyacetone kinase family protein
ORF01840 transcriptional regulator, TetR family, putative
ORF01842 hypothetical protein
ORF01843 dihydroxyacetone kinase family protein
ORF01844 dihydroxyacetone kinase family protein
ORF01847 conserved hypothetical protein
ORF01850 hypothetical protein
ORF01863 pyruvate phosphate dikinase (ppdK)
ORF01864 expressed protein of unknown function
ORF01865 CBS domain protein
ORF01866 3-hydroxyacyl-CoA dehydrogenase family protein, putative secreted protein
ORF01892 hypothetical protein
ORF01893 hypothetical protein
ORF01894 conserved hypothetical protein
ORF01895 hypothetical protein
ORF01896 hypothetical protein
ORF01897 hypothetical protein
ORF01898 hypothetical protein
ORF01899 hypothetical protein
ORF01903 conserved hypothetical protein
ORF01904 drug resistance transporter, EmrB/QacA family
ORF01905 hypothetical protein
ORF01922 conserved hypothetical protein
ORF01925 FMN-binding protein
ORF01934 hypothetical protein
ORF01936 polyprenyl synthetase family protein
ORF01939 cytochrome d ubiquinol oxidase, subunit II (cydB)
ORF01940 cytochrome d oxidase, subunit I (cydA)
ORF01941 pyridine nucleotide-disulphide oxidoreductase family protein
ORF01942 prenyltransferase, UbiA family
ORF01943 hypothetical protein
ORF01944 hypothetical protein
ORF01946 cyclopropane-fatty-acyl-phospholipid synthase (cfa)
ORF01951 conserved hypothetical protein
ORF01953 hypothetical protein
ORF01954 conserved hypothetical protein
ORF01984 hypothetical protein
ORF01988 hypothetical protein

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF01989 hypothetical protein
ORF01990 hypothetical protein
ORF01991 hypothetical protein
ORF02000 membrane protein, putative
ORF02001 transposase, IS30 family, putative
ORF02005 hypothetical protein
ORF02006 xylulose-5-phosphate/fructose-6-phosphate phosphoketolase (xfp)
ORF02009 conserved hypothetical protein
ORF02010 carbohydrate kinase, FGGY family
ORF02011 hypothetical protein
ORF02012 PTS system component, putative
ORF02015 glyoxylate reductase, NADH-dependent
ORF02016 hypothetical protein
ORF02025 hypothetical protein
ORF02026 hypothetical protein
ORF02030 glutamate--cysteine ligase-related protein
ORF02036 phosphinothricin N-acetyltransferase (pat)
ORF02039 conserved hypothetical protein
ORF02044 conserved hypothetical protein
ORF02045 conserved hypothetical protein
ORF02046 prophage LambdaSa2, lysin, putative
ORF02047 prophage LambdaSa2, holin, putative
ORF02048 conserved hypothetical protein
ORF02049 hypothetical protein
ORF02050 conserved domain protein
ORF02051 prophage LambdaSa2, PblB, putative
ORF02053 conserved hypothetical protein
ORF02056 conserved hypothetical protein
ORF02057 hypothetical protein
ORF02058 hypothetical protein
ORF02059 conserved hypothetical protein
ORF02060 conserved hypothetical protein
ORF02061 hypothetical protein
ORF02062 hypothetical protein
ORF02063 conserved domain protein
ORF02064 conserved domain protein
ORF02066 prophage LambdaSa2, protease, putative
ORF02067 conserved hypothetical protein
ORF02068 prophage LambdaSa2, terminase large subunit, putative
ORF02069 hypothetical protein
ORF02070 hypothetical protein
ORF02071 prophage LambdaSa2, site-specific recombinase, phage integrase family
ORF02072 conserved hypothetical protein
ORF02073 prophage LambdaSa2, transcriptional regulator, Cro/CI family
ORF02075 hypothetical protein
ORF02077 hypothetical protein
ORF02078 conserved hypothetical protein
ORF02079 conserved hypothetical protein
ORF02080 conserved hypothetical protein
ORF02081 hypothetical protein
ORF02084 prophage LambdaSa2, bacteriophage replication protein/hypothetical protein, truncation/fusion
ORF02085 hypothetical protein
ORF02087 hypothetical protein
ORF02088 conserved hypothetical protein

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF02089 prophage LambdaSa2, HNH endonuclease family protein
ORF02090 prophage LambdaSa2, antirepressor protein, putative
ORF02091 conserved domain protein
ORF02092 hypothetical protein
ORF02093 hypothetical protein
ORF02094 hypothetical protein
ORF02095 prophage LambdaSa2, repressor protein, putative
ORF02097 hypothetical protein
ORF02098 prophage LambdaSa2, site-specific recombinase, phage integrase family
ORF02100 hypothetical protein
ORF02102 hypothetical protein
ORF02103 microcin immunity protein MccF, putative
ORF02105 oxidoreductase, Gfo/Idh/MocA family
ORF02108 hypothetical protein
ORF02109 Cyclic nucleotide-binding domain protein
ORF02119 hypothetical protein
ORF02124 hypothetical protein
ORF02125 nitroreductase family protein
ORF02134 bacteriocin transport accessory protein, putative
ORF02148 neuraminidase-related protein
ORF02160 2`,3`-cyclic-nucleotide 2`-phosphodiesterase (cpdB)
ORF02163 conserved hypothetical protein
ORF02171 membrane protein, putative
ORF02172 hypothetical protein
ORF02173 membrane protein, putative
ORF02175 conserved hypothetical protein, truncation
ORF02181 phosphate transport system regulatory protein PhoU, putative
ORF02187 hypothetical protein
ORF02190 conserved hypothetical protein
ORF02191 hypothetical protein
ORF02194 acetyltransferase, GNAT family
ORF02196 hypothetical protein
ORF02198 acetyltransferase, GNAT family
ORF02201 membrane protein, putative
ORF02203 hypothetical protein
ORF02205 transcriptional regulator, Cro/CI family
ORF02206 conserved hypothetical protein
ORF02207 conserved hypothetical protein TIGR00730
ORF02208 hypothetical protein
ORF02209 site-specific recombinase, phage integrase family
ORF02210 conserved hypothetical protein
ORF02211 conserved hypothetical protein
ORF02212 hypothetical protein
ORF02213 hypothetical protein
ORF02214 transcriptional regulator, Cro/CI family
ORF02215 expressed protein of unknown function
ORF02216 site-specific recombinase, phage integrase family
ORF02217 conserved hypothetical protein
ORF02219 hypothetical protein
ORF02221 cell wall anchor protein-related protein
ORF02223 hypothetical protein
ORF02224 hypothetical protein
ORF02225 hypothetical protein
ORF02226 membrane protein, putative
ORF02227 conjugal transfer protein, interruption-C

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF02230 conserved hypothetical protein
ORF02231 conserved hypothetical protein
ORF02232 conserved hypothetical protein
ORF02235 hypothetical protein
ORF02236 conserved hypothetical protein
ORF02237 hypothetical protein
ORF02238 hypothetical protein
ORF02239 hypothetical protein
ORF02240 transcriptional regulator, Cro/CI family
ORF02241 hypothetical protein
ORF02242 transcriptional regulator, Cro/CI family
ORF02243 FtsK/SpoIIIE family protein
ORF02244 hypothetical protein
ORF02245 hypothetical protein
ORF02246 cell wall surface anchor family protein
ORF02247 transposase, ISL3 family
ORF02250 mercuric resistance operon regulatory protein MerR (merR)
ORF02251 Mn2+/Fe2+ transporter, NRAMP family
ORF02252 membrane protein, putative
ORF02253 ABC transporter, ATP-binding protein
ORF02254 conserved hypothetical protein
ORF02255 streptomycin resistance protein
ORF02257 hypothetical protein
ORF02258 hypothetical protein
ORF02259 conserved hypothetical protein
ORF02260 acetyltransferase, GNAT family
ORF02261 membrane protein, putative
ORF02263 hypothetical protein
ORF02264 transcriptional regulator, Cro/CI family
ORF02265 PAP2 family protein
ORF02266 conserved hypothetical protein FRAMESHIFT
ORF02267 conserved hypothetical protein TIGR00730
ORF02268 protease, putative
ORF02269 rhodanese family protein
ORF02271 hypothetical protein
ORF02274 conserved hypothetical protein
ORF02275 5-methyltetrahydrofolate--homocysteine methyltransferase, putative
ORF02277 conserved hypothetical protein
ORF02279 hypothetical protein
ORF02282 sensor histidine kinase
ORF02283 chromosome assembly-related protein
ORF02287 expressed protein of unknown function
ORF02291 pathogenicity protein, putative
ORF02308 hydrolase, haloacid dehalogenase-like family
ORF02314 conserved hypothetical protein
ORF02317 hypothetical protein
ORF02330 hypothetical protein
ORF02344 site-specific recombinase, phage integrase family
ORF02345 conserved hypothetical protein
ORF02346 conserved hypothetical protein
ORF02347 hypothetical protein
ORF02349 conserved hypothetical protein
ORF02350 hypothetical protein
ORF02351 transcriptional regulator, Cro/CI family
ORF02352 conserved domain protein

Table 11: GBS genes not shared with GAS or pneumococcus

ORFxxxxx Annotation
ORF02354 hypothetical protein
ORF02356 expressed putative secreted protein
ORF02362 sensor histidine kinase
ORF02363 response regulator
ORF02367 membrane protein, putative
ORF02368 conserved hypothetical protein
ORF02379 membrane protein, putative
ORF02395 transcriptional regulator, Cro/CI family
ORF02406 membrane protein, putative
ORF02416 diacylglycerol kinase catalytic domain protein, putative
ORF02418 hypothetical protein
ORF02422 hypothetical protein
ORF02425 conserved hypothetical protein
ORF03001 conserved hypothetical protein
ORF03004 conserved hypothetical protein
ORF03005 cylX protein
ORF03006 Tn916, hypothetical protein
ORF03007 Tn916, hypothetical protein
ORF03008 Tn916, hypothetical protein
ORF03009 Tn916, tetM leader peptide
ORF03010 Tn916, hypothetical protein
ORF03012 prophage LambdaSa2, HNH endonuclease family protein
ORF03013 conserved hypothetical protein
ORF03015 conjugal transfer protein, interruption-N

Table 12: GBS ORF's not shared with GAS, pneumococcus or any published genome

ORFxxxxx Annotation
ORF00035 membrane protein, putative
ORF00087 lipoprotein, putative
ORF00088 hypothetical protein
ORF00089 hypothetical protein
ORF00123 hypothetical protein
ORF00138 hypothetical protein
ORF00187 hypothetical protein
ORF00188 hypothetical protein
ORF00192 hypothetical protein
ORF00205 hypothetical protein
ORF00228 lipoprotein, putative
ORF00234 hypothetical protein
ORF00235 hypothetical protein
ORF00238 hypothetical protein
ORF00240 transcriptional regulator, Cro/CI family
ORF00241 hypothetical protein
ORF00242 conserved hypothetical protein
ORF00243 hypothetical protein
ORF00247 hypothetical protein
ORF00249 hypothetical protein
ORF00253 hypothetical protein
ORF00254 hypothetical protein
ORF00255 hypothetical protein
ORF00256 hypothetical protein
ORF00257 hypothetical protein
ORF00258 hypothetical protein
ORF00259 hypothetical protein
ORF00260 hypothetical protein
ORF00272 expressed putative lipoprotein
ORF00273 hypothetical protein
ORF00274 hypothetical protein
ORF00275 hypothetical protein
ORF00276 hypothetical protein
ORF00278 membrane protein, putative
ORF00285 lipoprotein, putative
ORF00292 hypothetical protein
ORF00294 expressed protein of unknown function
ORF00308 conserved hypothetical protein
ORF00332 hypothetical protein
ORF00340 hypothetical protein
ORF00384 hypothetical protein
ORF00402 membrane protein, putative
ORF00408 hypothetical protein
ORF00416 hypothetical protein
ORF00417 hypothetical protein
ORF00448 hypothetical protein
ORF00476 hypothetical protein
ORF00489 DNA-damage-inducible protein J, putative
ORF00490 hypothetical protein
ORF00491 lipoprotein, putative
ORF00497 conserved domain protein
ORF00510 bacteriocin transport accessory protein,putative
ORF00512 hypothetical protein
ORF00527 hypothetical protein
ORF00556 hypothetical protein

Table 12: GBS ORF's not shared with GAS, pneumococcus or any published genome

ORFxxxxx Annotation
ORF00575 hypothetical protein
ORF00599 hypothetical protein
ORF00618 hypothetical protein
ORF00620 hypothetical protein
ORF00623 hypothetical protein
ORF00626 prophage LambdaSa1, transcriptional regulator, Cro/CI family
ORF00628 hypothetical protein
ORF00630 hypothetical protein
ORF00632 hypothetical protein
ORF00635 hypothetical protein
ORF00636 hypothetical protein
ORF00637 hypothetical protein
ORF00642 conserved hypothetical protein
ORF00644 hypothetical protein
ORF00645 hypothetical protein
ORF00647 hypothetical protein
ORF00649 hypothetical protein
ORF00650 hypothetical protein
ORF00653 conserved hypothetical protein
ORF00657 conserved hypothetical protein, truncation
ORF00661 conserved hypothetical protein
ORF00673 hypothetical protein
ORF00674 hypothetical protein
ORF00675 conserved hypothetical protein
ORF00676 conserved hypothetical protein
ORF00682 hypothetical protein
ORF00685 conserved hypothetical protein
ORF00698 hypothetical protein
ORF00712 hypothetical protein
ORF00718 cell wall surface protein, interruption-N
ORF00723 hypothetical protein
ORF00735 expressed protein of unknown function
ORF00737 conserved hypothetical protein, degenerate
ORF00738 hypothetical protein
ORF00740 hypothetical protein
ORF00741 hypothetical protein
ORF00747 cylD protein (cylD)
ORF00753 cylE protein (cylE)
ORF00756 cylJ protein (cylJ)
ORF00757 cylK protein (cylK)
ORF00758 hypothetical protein
ORF00759 putative secreted protein
ORF00761 hypothetical protein
ORF00796 hydrolase, haloacid dehalogenase-like family
ORF00806 conserved hypothetical protein
ORF00822 ABC transporter, ATP-binding protein
ORF00827 hypothetical protein
ORF00872 cell wall surface anchor family protein
ORF00909 hypothetical protein
ORF00923 hypothetical protein
ORF00924 conserved hypothetical protein
ORF00942 expressed putative secreted protein
ORF00943 hypothetical protein
ORF00944 hypothetical protein
ORF01013 hypothetical protein

Table 12: GBS ORF's not shared with GAS, pneumococcus or any published genome

ORFxxxxx	Annotation
ORF01014	hypothetical protein
ORF01015	hypothetical protein
ORF01016	hypothetical protein
ORF01018	hypothetical protein
ORF01019	hypothetical protein
ORF01021	hypothetical protein
ORF01035	Tn916, excisionase
ORF01062	hypothetical protein
ORF01096	nisin-resistance protein, putative
ORF01145	lipoprotein, putative
ORF01146	conserved hypothetical protein, FRAMESHIFT
ORF01148	lipoprotein, putative
ORF01149	hypothetical protein
ORF01150	hypothetical protein
ORF01151	hypothetical protein
ORF01152	lipoprotein, putative
ORF01153	hypothetical protein
ORF01158	hypothetical protein
ORF01159	hypothetical protein
ORF01161	expressed conserved domain protein
ORF01162	conserved hypothetical protein
ORF01166	hypothetical protein
ORF01168	conserved hypothetical protein
ORF01169	hypothetical protein
ORF01174	conserved domain protein
ORF01175	hypothetical protein
ORF01186	cell wall surface anchor family protein, putative
ORF01187	hypothetical protein
ORF01204	hypothetical protein
ORF01215	hypothetical protein
ORF01258	hypothetical protein
ORF01262	conserved hypothetical protein, FRAMESHIFT
ORF01263	hypothetical protein
ORF01265	hypothetical protein
ORF01266	hypothetical protein
ORF01304	polysaccharide biosynthesis protein CpsK(V) (cpsK)
ORF01308	polysaccharide biosynthesis protein CpsM(V) (cpsM)
ORF01309	polysaccharide biosynthesis protein cpsH(V) (cpsH)
ORF01349	hypothetical protein
ORF01384	hypothetical protein
ORF01385	hypothetical protein
ORF01386	conserved hypothetical protein
ORF01392	hypothetical protein
ORF01395	conserved hypothetical protein
ORF01409	conserved hypothetical protein
ORF01410	hypothetical protein
ORF01417	hypothetical protein
ORF01418	hypothetical protein
ORF01420	hypothetical protein
ORF01423	conserved hypothetical protein
ORF01424	conserved hypothetical protein
ORF01425	conserved hypothetical protein
ORF01426	conserved hypothetical protein
ORF01427	hypothetical protein
ORF01431	hypothetical protein

Table 12: GBS ORF's not shared with GAS, pneumococcus or any published genome

ORFxxxxx Annotation
ORF01432 conserved domain protein
ORF01434 hypothetical protein
ORF01435 calcium-binding protein, putative
ORF01437 abortive infection protein AbiGI (abiGI)
ORF01438 abortive infection protein AbiGII (abiGII)
ORF01441 conserved hypothetical protein, degenerate
ORF01443 hypothetical protein
ORF01445 hypothetical protein
ORF01452 hypothetical protein
ORF01459 hypothetical protein
ORF01463 hypothetical protein
ORF01464 hypothetical protein
ORF01465 hypothetical protein
ORF01486 hypothetical protein
ORF01488 R5 protein
ORF01575 membrane protein, putative
ORF01581 lipoprotein, putative
ORF01601 hypothetical protein
ORF01611 hypothetical protein
ORF01638 conserved hypothetical protein
ORF01645 cell wall surface anchor family protein
ORF01660 membrane protein, putative
ORF01666 hypothetical protein
ORF01667 hypothetical protein
ORF01670 hypothetical protein
ORF01673 hypothetical protein
ORF01674 hypothetical protein
ORF01675 hypothetical protein
ORF01681 hypothetical protein
ORF01682 hypothetical protein
ORF01684 hypothetical protein
ORF01719 hypothetical protein
ORF01733 hypothetical protein
ORF01735 hypothetical protein
ORF01736 hypothetical protein
ORF01737 hypothetical protein
ORF01742 hypothetical protein
ORF01754 hypothetical protein
ORF01761 transposase, IS30 family, putative, truncation
ORF01807 hypothetical protein
ORF01836 hypothetical protein
ORF01838 hypothetical protein
ORF01842 hypothetical protein
ORF01850 hypothetical protein
ORF01892 hypothetical protein
ORF01893 hypothetical protein
ORF01895 hypothetical protein
ORF01896 hypothetical protein
ORF01897 hypothetical protein
ORF01898 hypothetical protein
ORF01899 hypothetical protein
ORF01905 hypothetical protein
ORF01934 hypothetical protein
ORF01943 hypothetical protein
ORF01944 hypothetical protein

Table 12: GBS ORF's not shared with GAS, pneumococcus or any published genome

ORFxxxxx Annotation
ORF01953 hypothetical protein
ORF01984 hypothetical protein
ORF01988 hypothetical protein
ORF01989 hypothetical protein
ORF02005 hypothetical protein
ORF02011 hypothetical protein
ORF02016 hypothetical protein
ORF02025 hypothetical protein
ORF02026 hypothetical protein
ORF02045 conserved hypothetical protein
ORF02047 prophage LambdaSa2, holin, putative
ORF02048 conserved hypothetical protein
ORF02049 hypothetical protein
ORF02050 conserved domain protein
ORF02053 conserved hypothetical protein
ORF02057 hypothetical protein
ORF02058 hypothetical protein
ORF02061 hypothetical protein
ORF02062 hypothetical protein
ORF02063 conserved domain protein
ORF02067 conserved hypothetical protein
ORF02069 hypothetical protein
ORF02070 hypothetical protein
ORF02072 conserved hypothetical protein
ORF02073 prophage LambdaSa2, transcriptional regulator, Cro/CI family
ORF02075 hypothetical protein
ORF02077 hypothetical protein
ORF02078 conserved hypothetical protein
ORF02081 hypothetical protein
ORF02085 hypothetical protein
ORF02087 hypothetical protein
ORF02088 conserved hypothetical protein
ORF02091 conserved domain protein
ORF02092 hypothetical protein
ORF02093 hypothetical protein
ORF02094 hypothetical protein
ORF02097 hypothetical protein
ORF02100 hypothetical protein
ORF02102 hypothetical protein
ORF02108 hypothetical protein
ORF02119 hypothetical protein
ORF02124 hypothetical protein
ORF02171 membrane protein, putative
ORF02172 hypothetical protein
ORF02173 membrane protein, putative
ORF02191 hypothetical protein
ORF02196 hypothetical protein
ORF02203 hypothetical protein
ORF02208 hypothetical protein
ORF02212 hypothetical protein
ORF02213 hypothetical protein
ORF02214 transcriptional regulator, Cro/CI family
ORF02215 expressed protein of unknown function
ORF02217 conserved hypothetical protein
ORF02219 hypothetical protein

Table 12: GBS ORF's not shared with GAS, pneumococcus or any published genome

ORFxxxxx Annotation
ORF02221 cell wall anchor protein-related protein
ORF02223 hypothetical protein
ORF02224 hypothetical protein
ORF02225 hypothetical protein
ORF02231 conserved hypothetical protein
ORF02235 hypothetical protein
ORF02236 conserved hypothetical protein
ORF02237 hypothetical protein
ORF02238 hypothetical protein
ORF02239 hypothetical protein
ORF02241 hypothetical protein
ORF02244 hypothetical protein
ORF02245 hypothetical protein
ORF02263 hypothetical protein
ORF02268 protease, putative
ORF02271 hypothetical protein
ORF02279 hypothetical protein
ORF02283 chromosome assembly-related protein
ORF02317 hypothetical protein
ORF02330 hypothetical protein
ORF02344 site-specific recombinase, phage integrase family
ORF02345 conserved hypothetical protein
ORF02347 hypothetical protein
ORF02349 conserved hypothetical protein
ORF02350 hypothetical protein
ORF02351 transcriptional regulator, Cro/CI family
ORF02354 hypothetical protein
ORF02356 expressed putative secreted protein
ORF02395 transcriptional regulator, Cro/CI family
ORF02418 hypothetical protein
ORF02422 hypothetical protein
ORF02425 conserved hypothetical protein
ORF03004 conserved hypothetical protein
ORF03005 cylX protein
ORF03006 Tn916, hypothetical protein
ORF03007 Tn916, hypothetical protein
ORF03008 Tn916, hypothetical protein
ORF03009 Tn916, tetM leader peptide
ORF03010 Tn916, hypothetical protein
ORF03015 conjugal transfer protein, interruption-N

Table 13: Comparative Sequences relating to SAG0466 (thiolase)

SEQ ID NO. 1301: SAG0466 FROM THE 2603V/R GBS STRAIN
CTCCTGCCCCTGCAATGGCAGTTAGACCCATAGGTTTATTTTTATATTTTAATGCCTGCATAAGATGAAGGATATTAATA
ATTCCTGAGCAGGCATAAGGGTGTCCGTAAGCTAATGTCCCTCCAAAAATATTGAATTTTTCTCTCTCTTCAGGATAATA
ATGATTAAATAGAGCATCAATCGCTGCAAATGGTTCATTCCATTCAATTGCATCATAATCCGATATTTTAGTATGAGTTT
CTGTTAATAGTTTTTCCGTAGCCGTGTGAACCAATTCTGGACTAAGCTTGGGATCTCCTGCTACTTCTACAATGTGAACA
ATCCGGAATTCTGTTTTCTGACTCTGAAGCGTTAGAAATGCAGCAGCATCGTGCATTAAACAAACATTTCCAATAGTGAG
CAAAGGTGAATTTTCCATCAATCTTGGTAATTTTTGAAAAAATGTTtCTTTTaGTTTTCTAACGCCTTGATCTCGCATCC
CTTCCATTGGTAAGATTACyTCTTCTAAATAGCCACCTTGTTTAGCTGTTAAGGCGCGTTTATGGCTCAAGAATGCCAAT
TTATCTAACATTTCTCTTCTAAAaCCATATTTTTGACAGACTCTCTGGGCCCCTTCTAACATTACAGTTTCAGCATAAGA
GTCAGGAGAAAACTGAGCAACTGTATATTCTCCGTTACGATTATCTTCTTTAGCATAACGTCTCATAGGTTGAAGAGAAC
TACTTTCAATCCCCCCAACAAGAACTTTTTCATTAATACCGGTACTGATTTTTAGATAACCAAAAAACAAGGCAGAACTT
GATGAAGCACACTGCATATCAATCGTTTGTACTGGAATATAGGATTCATAATCAGAAAAAAGAGTCATCAAACGACCAAT
ATTGCCCCCAGTACCAACTGTGTTCCCACAAATAATACTATCAATGTTAGATTCTGATTCTATTTTTTTTATTTGATTTA
AAAGGTGTGCTCCTAAAAGTTCTGGACGGTAAGTTTAAATTGCTT

SEQ ID NO. 1302: SAG0466 FROM THE M732 GBS TYPE III STRAIN
TCGGTATAAAAGGGAAGCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTCTTAAATCAAATAAAAAAAATA
GAATCAGAATCTAATATTGATAGTATTATTTGTGGGAACACAGTTGGTACTGGGGGCAATATTGGTCGTTTGATGACTCT
TTTTTCTGATTATGAATCCTATATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGTTCTGCCTTGTTTTTTG
GTTATCTAAAAATCAGTGCCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAAAGTAGTTCTCTTCAACCTATGAGA
CGTTACGCTAAAGAAGATAATCGTAACGGAGAATATACCGTTGCTCAGTTTTCTCCTGACTCTTATGCTGAAACTGTAAT
GTTAGAAGGGGCACAAAGAGTCTGTCAAAAATATGGTTTTAGAAGAGAAATGTTAGATAAATTGGCATTCTTGAGCCATA
AACGCGCCTTAACAGCTAAACAAGGTGGCTATTTAGAAGAGGTAATCTTACCAATGGAAGGGATGCGAGATCAAGGCGTT
AGAAAACTAAAAGAAGCATTTTTTCAAAAATTACCAAGATTGATGGAAAATTCACCTTTGCTCACTATTGGAAATGTTTG
TTTAATGCACGATGCTGCTGCATTTCTAACGCTTCAGAGTCAGAAAACAGAATTCCGGATTGTTCACATTGTAGAAGTAG
CAGGAGATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACAGAAACTCATACTAAAATATCG
GATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGATGCTTTATTTAATCATTATTATCCTGAAGAGAGAGA
AAAATTCAATATTTTTGGAGGGGCATTAGCTTACGGACACCCTTATGCCTGCTCAGGAATTA

SEQ ID NO. 1303: SAG0466 FROM THE 090 GBS TYPE Ia STRAIN
TTGTGGGAACACAGTTGGTACTGGGGGCAATATTGGTCGTTTGATGACTCTTTTTTCTGATTATGAATCCTATATTCCAG
TACAAACGATTGATATGCAGTGTGCTTCATCAAGTTCTGCCTTGTTTTTTGGTTATCTAAAAATCAGTGCCGGTATTAAT
GAAAAAGTTCTTGTTGGGGGGATTGAAAGTAGTTCTCTTCAACCTATGAGACGTTACGCTAAAGAAGATAATCGTAACGG
AGAATATACCGTTGCTCAGTTTTCTCCTGACTCTTAkGCTGAAACTGTAATGtTAGAAGGGGCACAAAGAGTCTGTCAAA
AATATGGTTTtAGAAGAGAAATGTTAGATAAATTGGCATTCTTGAGCCATAAACGCGCCTTAACAGCTAAACAAGGTGGC
TATTTAGAAGAGGTAATCTTACCAATGGAAGGGATGCGAGATCAAGGCGTTAGAAAACTAAAAGAAGCATTTTTTCAAAA
ATTACCAAGATTGATGGrAAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTCTwA
CGCTTCAGAGTCAGAAAACAGAATTCCGGATTGTTCACATTGTAGAAGTAGCAGGAGATCCCAAGCTTAGTCCAGAATTG
GTTCACACGGCTACGGAAAAACTATTAACAGAAACTCATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACC
ATTTGCAGCGATTGATGCTTTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGGCATTAG
CTTACGGACACCCTTATGCCTGCTCAGG

SEQ ID NO. 1304: SAG0466 FROM THE COH1 GBS TYPE Ia STRAIN
ATCGGTATAAAAGGGAAGCAATTTAAAATTACCGTCCAGAACTTTTAGGAGCACACCTCTTAAATCAAATAAAAAAAATA
GAATCAGAATCTAATATTGATAGTATTATTTGTGGGAACACAGTTGGTACTGGGGGCAATATTGGTCGTTTGATGACTCT
TTTTTCTGATTATGAATCCTATATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGTTCTGCCTTGTTTTTTG
GGTATCTAAAAA

SEQ ID NO. 1305 : SAG0466 FROM THE CJB GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)
TTTTCAAAAATTACCAAGATTGATGGAAAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTG
CATTTCTAACGCTTCAGAGTCAGAAAACAGAATTCCGGATTGTTCACATTGTAGAAGTAGCAGGAGATCCCAAGCTTAGT
CCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACAGAAACTCATACTAAAATATCGGATTATGATGCAATTGAATG
GAATGAACCATTTGCAGCGATTGATGCTTTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAG
GGGCATTAGCTTACGGACACCCTTAATGCCTGCTCAGGAATTATTAATATCC

SEQ ID NO. 1306: sag0466 FROM THE CJB110 GBS NONTYPEABLE STRAIN
GGTATAAAAGGGAAGCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTCTTAAATCAAATAAAAAAAATATA
ACCAGAATCTAACATTGATAGTATTATTTGTGGGAACACAGTTGGTACTGGGGGCAATATTGGTCGTTTGATGACTCTTT
TTTCTGATTATGAATCCTATATTC

Table 13: Comparative Sequences relating to SAG0466 (thiolase)

SEQ ID NO. 1307: SAG0466 FROM THE 1169NT1 GBS TYPE V STRAIN REVERSE COMPLEMENT
CAAGATTGATGGAAAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTCTAACGCTT
CAGAGTCAGAAAACAGAATTCCGGATTGTTCACATTGTAGAAGTAGCAGGAGATCCCAAGCTTAGTCCAGAATTGGTTCA
CACGGCTACGGAAAAACTATTAACAGAAACTCATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTG
CAGCGATTGATGCTTTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGGCATTAGCTTAC
GGACACCCTTATGCCTGCTCAGGAATTATTAATATCCTTCATCTTATGCAGGCATTAAAATATAAAAATAAACCTATGGG
CCTAACTGCCATTGCAGGGGCA

SEQ ID NO. 1308: SAG0466 FROM THE 18RS21 GBS TYPE II STRAIN
CCTTAACAGTTAAACAAGGTGGCTATTTAGAAGAGGTAATCTTACCAATGGAAGGGATGCGAGATCAAGGCGTTAGAAAA
CTAAAAGAAACATTTTTTCAAAAATTACCAAGATTGATGGAAAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAAT
GCACGATGCTGCTGCATTTCTAACGCTTCAGAGTCAGAAAACAGAATTCCGGATTGTTCACATTGTAGAAGTAGCAGGAG
ATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACAGAAACTCATACTAAAATATCGGATTAT
GATGCAATTGAATGGAATGAACCATTTGCAGCGATTGATGCTCTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATT
CAATATTTTTGGAGGGACATTAGCTTACGGACACCCTTATGCCTGCTCAGGAATTATTAATATCCTTCATCTTATGCAGG
CATTAAAATATAAAAATAAACCTATGGGTCTAACTGCCATTGCAGGGGCAG

SEQ ID NO. 1309: SAG0466 FROM THE 18RS21 GBS TYPE II STRAIN
TCGGTATAAAAGGGAAGCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTTTTAAATCAAATAAAAAAAATA
GAATCAGAATCTAACATTGATAGTATTATTTGTGGAACACAGTTGGTACTGGGGGCAATATTGGTCGTTTGATGACTCT
TTTTTCTGATTATGAATCCTATATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGTTCTGCCTTGTTTTTTG
GTTATCTAAAAATCAGTACCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAAAGTAGTTCTCTTCAACCTATGAGA
CGTTATGCTAAAGAAGATAATCGTAACGGAGAATATACAGTTGCTCAGTTTTCTCCTGACTCTTATGCTGAAACTGTAAT
GTTAGAAGGGGCCCAGAGAGTCTGTCAAAAATATGGTTTTAGAAGAGAAATGTTAGATAAATTGGCATTCTTGAGCCATA
AACGCGCCTTAACAGCTAAACA

SEQ ID NO. 1310: SAG0466 FROM THE H36b GBS TYPE Ib STRAIN
TTTGGGCTACGAACACCTATCGGTATAAAAGGGAAGCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTTTT
AAATCAAATAAAAAAAATAGAATCAGAATCTAACATTGATAGTATTATTTGTGGAACACAGTTGGTACTGGGGGCAATA
TTGGTCGTTTGATGACTCTTTTTTCTGATTATGAATCCTATATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCA
AGTTCTGCCTTGTTTTTTGGTTATCTAAAAATCAGTACCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAAAGTAG
TTCTCTTCAACCTATGAGACGTTATGCTAAAGAAGATAATCGTAACGGAGAATATACAGTTGCTCAGTTTTCTCCTGACT
CTTATGCTGAAACTGTAATGTTAGAAGGGGCCC

SEQ ID NO. 1311: SAG0466 FROM THE H36b GBS TYPE Ib STRAIN (REVERSE COMPLEMENT)
GAAAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTCTAACGCTTCAGAGTCAGAA
AACAGAATTCCGGATTGTTCACATTGTAGAAGTAGCAGGAGATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGG
AAAAACTATTAACAGAAACTCATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGAT
GCTCTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGACATTAGCTTACGGACACCCTTA
TGCCTGCTCAGGAATTATTAATATCCTTCATCTTATGCAGGCATTAAAATATAAAAATAAACCTATGGGTCTAACTGCCA
TTGCAGGGGCAGGA

SEQ ID NO. 1312: SAG0466 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
CCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTCTAACGCTTCAGAGTCAGAAAACAGAATT
CCGGATTGTTCACATTGTAGAAGTAGCAGGAGATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTAT
TAACAGAAACTCATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGATGCTTTATTT
AATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGGCATTAGCTTACGGACACCCTTATGCCTGCTC
AGGAATTATTAATATCCTTCATCTTATGCAGGCATTAAAATATAAAAATAAACCTATGGGTTCTAACTGC

SEQ ID NO. 1313: SAG0466 FROM THE M781 GBS TYPE III STRAIN
GCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTCTTAAATCAAATAAAAAAAATAGAATCAGAATCTAATA
TTGATAGTATTATTTGTGGAACACAGTTGGTACTGGGGGCAATATTGGTCGTTTGATGACTCTTTTTTCTGATTATGAA
TCCTATATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGTTCTGCCTTGTTTTTTGGTTATCTAAAAATCAG
TGCCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAAAGTAGTTCTCTTCAACCTATGAGACGTTACGCTAAAGAAG
ATAATCGTAACGGAGAATATACCGTTGCTCAGTTTTCTCCTGACTCTTATGCTGAAACTGTAATGTTAGA

SEQ ID NO 1314: SAG0466 FROM THE A909 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
CCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTCTAACGCTTCAGAGTCAGAAAACAGAATT
CCGGATTGTTCACATTGTAGAAGTAGCAGGAGATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTAT
TAACAGAAACTCATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGATGCTCTATTT
AATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGACATTAGCTTACGGACACCCTTATGCCTGCTC
AGGAATTATTAATATCCTTCATCTTATGCAGGCATTAAAATATAAAAATAAACCTATGGGTCTAACTGCCATTGCAGGGG
C

Table 13: Comparative Sequences relating to SAG0466 (thiolase)

SEQ ID NO. 1315: SAG0466 FROM THE JM9130013 GBS TYPE VIII STRAIN REVERSE COMPLEMENT

GCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTCTAACGCTTCAGAGTCAGAAAACAGAATTCCGGA
TTGTTCACATTGTAGAAGTAGCAGGAGATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACA
GAAACTCATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGATGCTCTATTTAATCA
TTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTGGAGGGCATTAGCTTACGGACACCCTTATGCCTGCTCAGGAA
TTATTAATATCCTTCATCTTATGCAGGCATTAAAATATAAAAATAAACCTATGGGTCTAACTGCCATTGCAGGGGCAGGA

SEQ ID NO. 1316: SAG0466 FROM THE JM9130013 GBS TYPE VIII STRAIN

TTTGGGCTACGAACACCTATCGGTATAAAAGGGAAGCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTTTT
AAATCAAATAAAAAAAATAGAATCAGAATCTAACATTGATAGTATTATTTGTGGGAACACAGTTGGTACTGGGGGCAATA
TTGGTCGTTTGATGACTCTTTTTTCTGATTATGAATCCTATATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCA
AGTTCTGCCTTGTTTTTTGGTTATCTAAAAATCAGTACCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAAAGTAG
TTCTCTTCAACCTATGAGACGTTATGCTAAAGAAGATAATCGTAACGGAGAATATA

```
SEQ1301      --------CTCCTGCCCCTGCAATGGCAGTTAGACCCATAGGTTTATTTTTATATTTTA
SEQ1302      -----------------------------------------------------------
SEQ1303      -----------------------------------------------------------
SEQ1304      -----------------------------------------------------------
SEQ1305      -----------------------------------------------------------
SEQ1306      -----------------------------------------------------------
SEQ1307      -----------------------------------------------------------
SEQ1308      CTTAACAGTTAAACAAGGTGGCTATTTAGAAGAGGTAATCTTACCAATGGAAGGGATGC
SEQ1309      -----------------------------------------------------------
SEQ1310      -----------------------------------------------------------
SEQ1311      -----------------------------------------------------------
SEQ1312      -----------------------------------------------------------
SEQ1313      -----------------------------------------------------------
SEQ1314      -----------------------------------------------------------
SEQ1315      -----------------------------------------------------------
SEQ1316      -----------------------------------------------------------

SEQ1301      TGCCTGCATAAGATGAAGGATATTAATAATTCCTGAGCAGGCATAAGGGTGTCCGTAAG
SEQ1302      -------------------------------------------------TCGGTATAAA
SEQ1303      -----------------------------------------------------------
SEQ1304      ------------------------------------------------ATCGGTATAAA
SEQ1305      ---------------------------------TTTTCAAAAATTACCAAGATTGATGG
SEQ1306      ---------------------------------------------------GGTATAAA
SEQ1307      -----------------------------------------------CAAGATTGATGG
SEQ1308      AGATCAAGGCGTTAGAAAACTAAAAGAAACATTTTTTCAAAAATTACCAAGATTGATGG
SEQ1309      -------------------------------------------------TCGGTATAAA
SEQ1310      ------------------------------TTTGGGCTACGAACACCTATCGGTATAAA
SEQ1311      ----------------------------------------------------------G
SEQ1312      -----------------------------------------------------------
SEQ1313      -----------------------------------------------------------
SEQ1314      -----------------------------------------------------------
SEQ1315      -----------------------------------------------------------
SEQ1316      ------------------------------TTTGGGCTACGAACACCTATCGGTATAAA

SEQ1301      TAATGTCCCTCCAAA-AATATTGAATTTTTCTCTCTC-TTCAGGATAATAATGATTAAA
SEQ1302      GGGAAGCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTCTTAAATCAAAT
SEQ1303      -----------------------------------------------------------
SEQ1304      GGGAAGCAATTTAAA-ATTACCGTCCAGAACTTTTAGGAGCACACCTCTTAAATCAAAT
SEQ1305      AAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTC
SEQ1306      GGGAAGCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTCTTAAATCAAAT
SEQ1307      AAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTC
SEQ1308      AAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTC
SEQ1309      GGGAAGCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTTTTAAATCAAAT
SEQ1310      GGGAAGCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTTTTAAATCAAAT
SEQ1311      AAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTC
SEQ1312      -------CCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTC
SEQ1313      -----GCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTCTTAAATCAAAT
SEQ1314      -------CCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTC
SEQ1315      ------------GCTCACTATTGGAAATGTTTGTTTAATGCACGATGCTGCTGCATTTC
```

Table 13: Comparative Sequences relating to SAG0466 (thiolase)

SEQ1316	GGGAAGCAATTTAAACATTACCGTCCAGAACTTTTAGGAGCACACCTTTTAAATCAAAT

SEQ1301	AGAGCATCAATCGCTGCAAATGGTTCATTCC-ATTCAATTGCATCATAATCCGATATTT
SEQ1302	AAAAAAATAGAATCAGAATCTAAT--ATT---GATAGTATTATTTGTGGGAACA-CAGT
SEQ1303	---TTGTGGGAACA-CAGT
SEQ1304	AAAAAAATAGAATCAGAATCTAAT--ATT---GATAGTATTATTTGTGGGAACA-CAGT
SEQ1305	AACGCTTCAGAGTCAGAAAACAGA--ATTCCGGATTGTTCACATTGTAGAAGTAGCAGG
SEQ1306	AAAAAAATATAACCAGAATCTAAC--ATT---GATAGTATTATTTGTGGGAACA-CAGT
SEQ1307	AACGCTTCAGAGTCAGAAAACAGA--ATTCCGGATTGTTCACATTGTAGAAGTAGCAGG
SEQ1308	AACGCTTCAGAGTCAGAAAACAGA--ATTCCGGATTGTTCACATTGTAGAAGTAGCAGG
SEQ1309	AAAAAAATAGAATCAGAATCTAAC--ATT---GATAGTATTATTTGTGGGAACA-CAGT
SEQ1310	AAAAAAATAGAATCAGAATCTAAC--ATT---GATAGTATTATTTGTGGGAACA-CAGT
SEQ1311	AACGCTTCAGAGTCAGAAAACAGA--ATTCCGGATTGTTCACATTGTAGAAGTAGCAGG
SEQ1312	AACGCTTCAGAGTCAGAAAACAGA--ATTCCGGATTGTTCACATTGTAGAAGTAGCAGG
SEQ1313	AAAAAAATAGAATCAGAATCTAAT--ATT---GATAGTATTATTTGTGGGAACA-CAGT
SEQ1314	AACGCTTCAGAGTCAGAAAACAGA--ATTCCGGATTGTTCACATTGTAGAAGTAGCAGG
SEQ1315	AACGCTTCAGAGTCAGAAAACAGA--ATTCCGGATTGTTCACATTGTAGAAGTAGCAGG
SEQ1316	AAAAAAATAGAATCAGAATCTAAC--ATT---GATAGTATTATTTGTGGGAACA-CAGT

SEQ1301	AGTATGAGTTTCTGTTAATAGTTTTTCCGTAGCCGTGTGAACCAATTCTGGACTAAGCT
SEQ1302	GGTACTGGGGGCAATATTGG-TCGTTTGATGACTCTTTTTTCTGATTATGAATCCTA--
SEQ1303	GGTACTGGGGGCAATATTGG-TCGTTTGATGACTCTTTTTTCTGATTATGAATCCTA--
SEQ1304	GGTACTGGGGGCAATATTGG-TCGTTTGATGACTCTTTTTTCTGATTATGAATCCTA--
SEQ1305	GATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACAGAAAC
SEQ1306	GGTACTGGGGGCAATATTGG-TCGTTTGATGACTCTTTTTTCTGATTATGAATCCTA--
SEQ1307	GATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACAGAAAC
SEQ1308	GATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACAGAAAC
SEQ1309	GGTACTGGGGGCAATATTGG-TCGTTTGATGACTCTTTTTTCTGATTATGAATCCTA--
SEQ1310	GGTACTGGGGGCAATATTGG-TCGTTTGATGACTCTTTTTTCTGATTATGAATCCTA--
SEQ1311	GATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACAGAAAC
SEQ1312	GATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACAGAAAC
SEQ1313	GGTACTGGGGGCAATATTGG-TCGTTTGATGACTCTTTTTTCTGATTATGAATCCTA--
SEQ1314	GATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACAGAAAC
SEQ1315	GATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTATTAACAGAAAC
SEQ1316	GGTACTGGGGGCAATATTGG-TCGTTTGATGACTCTTTTTTCTGATTATGAATCCTA--

SEQ1301	GGGATCTCCTGCTACTTCTACAATGTGAACAATCCGGA-ATTCTGTTTTCTGACTCTGA
SEQ1302	TATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGT--TCTGCCTTGTTTTT
SEQ1303	TATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGT--TCTGCCTTGTTTTT
SEQ1304	TATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGT--TCTGCCTTGTTTTT
SEQ1305	CATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGA
SEQ1306	TATTC--
SEQ1307	CATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGA
SEQ1308	CATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGA
SEQ1309	TATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGT--TCTGCCTTGTTTTT
SEQ1310	TATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGT--TCTGCCTTGTTTTT
SEQ1311	CATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGA
SEQ1312	CATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGA
SEQ1313	TATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGT--TCTGCCTTGTTTTT
SEQ1314	CATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGA
SEQ1315	CATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTGCAGCGATTGA
SEQ1316	TATTCCAGTACAAACGATTGATATGCAGTGTGCTTCATCAAGT--TCTGCCTTGTTTTT

Table 13: Comparative Sequences relating to SAG0466 (thiolase)

```
SEQ1301     GCGTTAGAAATGCAGCAGCATCGTGCATTAAACAAACATTTC--CAATAGTGAGCAAAG
SEQ1302     GGT-TATCTAAAAATCAGTG-CCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAA
SEQ1303     GGT-TATCTAAAAATCAGTG-CCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAA
SEQ1304     GGG-TATCTAAAAA---------------------------------------------
SEQ1305     GCTTTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGGC
SEQ1306     -----------------------------------------------------------
SEQ1307     GCTTTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGGC
SEQ1308     GCTCTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGAC
SEQ1309     GGT-TATCTAAAAATCAGTA-CCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAA
SEQ1310     GGT-TATCTAAAAATCAGTA-CCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAA
SEQ1311     GCTCTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGAC
SEQ1312     GCTTTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGGC
SEQ1313     GGT-TATCTAAAAATCAGTG-CCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAA
SEQ1314     GCTCTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGAC
SEQ1315     GCTCTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTTTTGGAGGGGC
SEQ1316     GGT-TATCTAAAAATCAGTA-CCGGTATTAATGAAAAAGTTCTTGTTGGGGGGATTGAA

SEQ1301     TGAATTTTCCATCAATCTTGG--TAATTTTTGAAAAAATGTTTCTTTTAGTTTTCTAAC
SEQ1302     GTAGTTCTCTTCAACCTATGAGACGTTACGCTAAAGAAGATAATCGTAACGGAGAATAT
SEQ1303     GTAGTTCTCTTCAACCTATGAGACGTTACGCTAAAGAAGATAATCGTAACGGAGAATAT
SEQ1304     -----------------------------------------------------------
SEQ1305     TTAGCTTACGGACACCCTTAA--TGCCTGCTCAGGAATTATTAATATCC----------
SEQ1306     -----------------------------------------------------------
SEQ1307     TTAGCTTACGGACACCCTTA---TGCCTGCTCAGGAATTATTAATATCCTTCATCTTAT
SEQ1308     TTAGCTTACGGACACCCTTA---TGCCTGCTCAGGAATTATTAATATCCTTCATCTTAT
SEQ1309     GTAGTTCTCTTCAACCTATGAGACGTTATGCTAAAGAAGATAATCGTAACGGAGAATAT
SEQ1310     GTAGTTCTCTTCAACCTATGAGACGTTATGCTAAAGAAGATAATCGTAACGGAGAATAT
SEQ1311     TTAGCTTACGGACACCCTTA---TGCCTGCTCAGGAATTATTAATATCCTTCATCTTAT
SEQ1312     TTAGCTTACGGACACCCTTA---TGCCTGCTCAGGAATTATTAATATCCTTCATCTTAT
SEQ1313     GTAGTTCTCTTCAACCTATGAGACGTTACGCTAAAGAAGATAATCGTAACGGAGAATAT
SEQ1314     TTAGCTTACGGACACCCTTA---TGCCTGCTCAGGAATTATTAATATCCTTCATCTTAT
SEQ1315     TTAGCTTACGGACACCCTTA---TGCCTGCTCAGGAATTATTAATATCCTTCATCTTAT
SEQ1316     GTAGTTCTCTTCAACCTATGAGACGTTATGCTAAAGAAGATAATCGTAACGGAGAATAT

SEQ1301     CCTTGATCTCGCATCCCTTCCATTGGTAAGATTACYTCTTCTAAATAGCCACCTTGTTT
SEQ1302     CCGTTGCTCAGTTTTCTCCTGACTCTTATGCTG--AAACTGTAATGTTAGAAGGGGCAC
SEQ1303     CCGTTGCTCAGTTTTCTCCTGACTCTTAKGCTG--AAACTGTAATGTTAGAAGGGGCAC
SEQ1304     -----------------------------------------------------------
SEQ1305     -----------------------------------------------------------
SEQ1306     -----------------------------------------------------------
SEQ1307     CAGGCATTAAAATATAAAAATAAACCTATGGGC-CTAACTGCCATTGCAGGGGCA----
SEQ1308     CAGGCATTAAAATATAAAAATAAACCTATGGGT-CTAACTGCCATTGCAGGGGCAG---
SEQ1309     CAGTTGCTCAGTTTTCTCCTGACTCTTATGCTG--AAACTGTAATGTTAGAAGGGGCCC
SEQ1310     CAGTTGCTCAGTTTTCTCCTGACTCTTATGCTG--AAACTGTAATGTTAGAAGGGGCCC
SEQ1311     CAGGCATTAAAATATAAAAATAAACCTATGGGT-CTAACTGCCATTGCAGGGGCAGGA-
SEQ1312     CAGGCATTAAAATATAAAAATAAACCTATGGGTTCTAACTGC-----------------
SEQ1313     CCGTTGCTCAGTTTTCTCCTGACTCTTATGCTG--AAACTGTAATGTTAGA--------
SEQ1314     CAGGCATTAAAATATAAAAATAAACCTATGGGT-CTAACTGCCATTGCAGGGGC-----
SEQ1315     CAGGCATTAAAATATAAAAATAAACCTATGGGT-CTAACTGCCATTGCAGGGGCAGGA-
SEQ1316     TABCMARATVSTNCSRATNGTSAGTHAS-------------------------------
```

Table 13: Comparative Sequences relating to SAG0466 (thiolase)

```
SEQ1301     GCTGTTAAGGCGCGTTTATGGCTCAAGAATGCCAATTTATCTAACATTTCTCTTCTAAA
SEQ1302     AAGAGTCTGTCAAAAATATGGTTTTAGAAGAGAAATGTTAGATAAATTGGCATTCTTGA
SEQ1303     AAGAGTCTGTCAAAAATATGGTTTTAGAAGAGAAATGTTAGATAAATTGGCATTCTTGA
SEQ1304     -----------------------------------------------------------
SEQ1305     -----------------------------------------------------------
SEQ1306     -----------------------------------------------------------
SEQ1307     -----------------------------------------------------------
SEQ1308     -----------------------------------------------------------
SEQ1309     GAGAGTCTGTCAAAAATATGGTTTTAGAAGAGAAATGTTAGATAAATTGGCATTCTTGA
SEQ1310     -----------------------------------------------------------
SEQ1311     -----------------------------------------------------------
SEQ1312     -----------------------------------------------------------
SEQ1313     -----------------------------------------------------------
SEQ1314     -----------------------------------------------------------
SEQ1315     -----------------------------------------------------------
SEQ1316     -----------------------------------------------------------

SEQ1301     CCATATTTTTGACAGACTCTCTGGGCCCCTT--CTAACATTACAGTTTCAGCATAAGAG
SEQ1302     CCATAAACGCGCCTTAACAGCTAAACAAGGTGGCTATTTAGAAGAGGTAATCTTACCAA
SEQ1303     CCATAAACGCGCCTTAACAGCTAAACAAGGTGGCTATTTAGAAGAGGTAATCTTACCAA
SEQ1304     -----------------------------------------------------------
SEQ1305     -----------------------------------------------------------
SEQ1306     -----------------------------------------------------------
SEQ1307     -----------------------------------------------------------
SEQ1308     -----------------------------------------------------------
SEQ1309     CCATAAACGCGCCTTAACAGCTAAACA--------------------------------
SEQ1310     -----------------------------------------------------------
SEQ1311     -----------------------------------------------------------
SEQ1312     -----------------------------------------------------------
SEQ1313     -----------------------------------------------------------
SEQ1314     -----------------------------------------------------------
SEQ1315     -----------------------------------------------------------
SEQ1316     -----------------------------------------------------------

SEQ1301     CAGGAGAAAACTGAGCAACTGTATATTCTCCGTTACGATTATCTTCTTTAGCATAACGT
SEQ1302     GGAAGGGATGCGAGATCAAGGCGTTAGAAAACTAAAAGAAGCATTTTTTCAAAAATTAC
SEQ1303     GGAAGGGATGCGAGATCAAGGCGTTAGAAAACTAAAAGAAGCATTTTTTCAAAAATTAC
SEQ1304     -----------------------------------------------------------
SEQ1305     -----------------------------------------------------------
SEQ1306     -----------------------------------------------------------
SEQ1307     -----------------------------------------------------------
SEQ1308     -----------------------------------------------------------
SEQ1309     -----------------------------------------------------------
SEQ1310     -----------------------------------------------------------
SEQ1311     -----------------------------------------------------------
SEQ1312     -----------------------------------------------------------
SEQ1313     -----------------------------------------------------------
SEQ1314     -----------------------------------------------------------
SEQ1315     -----------------------------------------------------------
SEQ1316     -----------------------------------------------------------
```

Table 13: Comparative Sequences relating to SAG0466 (thiolase)

```
SEQ1301       TCATAGGTTGAAGAGAACTACTTTCAATCCCCCCAACAAGAACTTTTTCATTAATACCG
SEQ1302       AAGATTGATGGAAAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATG
SEQ1303       AAGATTGATGGRAAATTCACCTTTGCTCACTATTGGAAATGTTTGTTTAATGCACGATG
SEQ1304       ------------------------------------------------------------
SEQ1305       ------------------------------------------------------------
SEQ1306       ------------------------------------------------------------
SEQ1307       ------------------------------------------------------------
SEQ1308       ------------------------------------------------------------
SEQ1309       ------------------------------------------------------------
SEQ1310       ------------------------------------------------------------
SEQ1311       ------------------------------------------------------------
SEQ1312       ------------------------------------------------------------
SEQ1313       ------------------------------------------------------------
SEQ1314       ------------------------------------------------------------
SEQ1315       ------------------------------------------------------------
SEQ1316       ------------------------------------------------------------

SEQ1301       TACTGATTTTTAGATAACCAAAAAAC--AAGGCAGAACTTGATGAAGCACACTGCATAT
SEQ1302       TGCTGCATTTCTAACGCTTCAGAGTCAGAAAACAGAATTCCGGATTGTTCACATTGTAG
SEQ1303       TGCTGCATTTCTWACGCTTCAGAGTCAGAAAACAGAATTCCGGATTGTTCACATTGTAG
SEQ1304       ------------------------------------------------------------
SEQ1305       ------------------------------------------------------------
SEQ1306       ------------------------------------------------------------
SEQ1307       ------------------------------------------------------------
SEQ1308       ------------------------------------------------------------
SEQ1309       ------------------------------------------------------------
SEQ1310       ------------------------------------------------------------
SEQ1311       ------------------------------------------------------------
SEQ1312       ------------------------------------------------------------
SEQ1313       ------------------------------------------------------------
SEQ1314       ------------------------------------------------------------
SEQ1315       ------------------------------------------------------------
SEQ1316       ------------------------------------------------------------

SEQ1301       AATCGTTTGTACTGGAATATAGGATTCATAATCAGAAAAAAGAGTCATCAAACGACCAA
SEQ1302       AGTAGCAGGAGATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTAT
SEQ1303       AGTAGCAGGAGATCCCAAGCTTAGTCCAGAATTGGTTCACACGGCTACGGAAAAACTAT
SEQ1304       ------------------------------------------------------------
SEQ1305       ------------------------------------------------------------
SEQ1306       ------------------------------------------------------------
SEQ1307       ------------------------------------------------------------
SEQ1308       ------------------------------------------------------------
SEQ1309       ------------------------------------------------------------
SEQ1310       ------------------------------------------------------------
SEQ1311       ------------------------------------------------------------
SEQ1312       ------------------------------------------------------------
SEQ1313       ------------------------------------------------------------
SEQ1314       ------------------------------------------------------------
SEQ1315       ------------------------------------------------------------
SEQ1316       ------------------------------------------------------------
```

Table 13: Comparative Sequences relating to SAG0466 (thiolase)

```
SEQ1301        ATTGCCCCCAGTACCAACTGTGTTCCCACAAATAATACTATCAATGTTAGATTCTGATT
SEQ1302        AACAGAAACTCATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTG
SEQ1303        AACAGAAACTCATACTAAAATATCGGATTATGATGCAATTGAATGGAATGAACCATTTG
SEQ1304        -----------------------------------------------------------
SEQ1305        -----------------------------------------------------------
SEQ1306        -----------------------------------------------------------
SEQ1307        -----------------------------------------------------------
SEQ1308        -----------------------------------------------------------
SEQ1309        -----------------------------------------------------------
SEQ1310        -----------------------------------------------------------
SEQ1311        -----------------------------------------------------------
SEQ1312        -----------------------------------------------------------
SEQ1313        -----------------------------------------------------------
SEQ1314        -----------------------------------------------------------
SEQ1315        -----------------------------------------------------------
SEQ1316        -----------------------------------------------------------

SEQ1301        TATTTTTTTTATTTGATTTAAAAGGTGTGCTCCTAAAAGTTCTGGACGGTAAGTTTAAA
SEQ1302        AGCGATTGATGCTTTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTT
SEQ1303        AGCGATTGATGCTTTATTTAATCATTATTATCCTGAAGAGAGAGAAAAATTCAATATTT
SEQ1304        -----------------------------------------------------------
SEQ1305        -----------------------------------------------------------
SEQ1306        -----------------------------------------------------------
SEQ1307        -----------------------------------------------------------
SEQ1308        -----------------------------------------------------------
SEQ1309        -----------------------------------------------------------
SEQ1310        -----------------------------------------------------------
SEQ1311        -----------------------------------------------------------
SEQ1312        -----------------------------------------------------------
SEQ1313        -----------------------------------------------------------
SEQ1314        -----------------------------------------------------------
SEQ1315        -----------------------------------------------------------
SEQ1316        -----------------------------------------------------------

SEQ1301        TGCTT------------------------------------------
SEQ1302        TGGAGGGGCATTAGCTTACGGACACCCTTATGCCTGCTCAGGAATTA
SEQ1303        TGGAGGGGCATTAGCTTACGGACACCCTTATGCCTGCTCAGG-----
SEQ1304        -----------------------------------------------
SEQ1305        -----------------------------------------------
SEQ1306        -----------------------------------------------
SEQ1307        -----------------------------------------------
SEQ1308        -----------------------------------------------
SEQ1309        -----------------------------------------------
SEQ1310        -----------------------------------------------
SEQ1311        -----------------------------------------------
SEQ1312        -----------------------------------------------
SEQ1313        -----------------------------------------------
SEQ1314        -----------------------------------------------
SEQ1315        -----------------------------------------------
SEQ1316        -----------------------------------------------
```

Table 14: Comparative Sequences relating to SAG0471 (glucokinase)

SEQ ID NO. 1401: SAG0471 FROM THE 18RS21 GBS TYPE II STRAIN
TTAAATTTGGTATCTTGACGCTTGAGGGAGAAGTACAAGAAAAATGGGCAATTGAGACCAATACTTTAGAAAACGGAAGACATATCGTT
TCTGATATCGTTGAATCTCTCAAACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGGTATCGGTATGGGTTCTCCAGG
AGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGCTTTTAATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCAGTTATTGAAAAAG
AAGTTGGAATTCCATTTTTTATTGATAACGATGCTAATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGCTGGTGCCAATAATCCCGAC
GTTGTTTTCGTAACCCTCGGAACAGGAGTAGGTGGAGGTGTTATCGCAGATGGTAACCTCATCCATGGTGTTGCAGGAGCAGGTGGAGA
AATTGGGCATATGATTGTTGATCCAGAAAATGGATTTACGTGCACATGTGGTAACAAAGGCTGCCTTGAGACAGTTGCATCAGCGACAG
GTGTTGTTAGAGTAGCACGTCAACTCGCAGAACAATATGAGGGTTCGTCTGCCATTAAAGCAGCGATTGACACCGGTGATACTGTTACA
AGTAAAGATATTTTTATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTATCACGTTACCTTGGACTGGCAGC
AGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATTGGTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCGTTG
AGAAATACTTTGTCACATTTGCTTTCCCACAAGTTAAAAAGTCAACTAAAATTAAGAT

SEQ ID NO. 1402: SAG0471 FROM THE 090 GBS TYPE Ia STRAIN
CGTTTCTGATATCGTTGAATCTCTCAAACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGGTATCGGTATGGGTTCTC
CAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGCTTTTAATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCGGTTATTGAA
AAAGAAGTTGGAATTCCATTTTTTATTGATAACGATGCTAATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGCTGGTGCCAATAATCC
CGATGTTGTTTTCGTAACCCTCGGAACAGGAGTAGGTGGAGGTGTTATCGCAGATGGTAACCTCATCCATGGTGTTGCAGGAGCAGGTG
GAGAAATTGGGCATATGATTGTTGATCCAGAKAATGGATTTACGTGCACATGTGGTAACAAAGGCTGTCTTGAGACAGTTGCATCAGCG
ACAGGTGTTGTTAGAGTAGCACGTCAACTCGCAGAACAATATGAAGGTTCGTCTGCCATTAAAGCAGCGATTGACAACGGTGATACTGT
TACAAGTAAAGATATTTTTATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTATCACGTTACCTTGGACTGG
CAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATTGGTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGC
GTTGAGAAATACTTTGTCACATTTG

SEQ ID NO. 1403: SAG0471 FROM THE COH1 GBS TYPE Ia STRAIN
ACAAGAAAAAATGGGCAATTGAGACCAATACTTTAGAAAACGGAAGACATATCGTTTCTGATATCGTTGAATCTCTCAAACATCGTTTGA
GCCTCTATGGATTAACAAAAGATGACTTTCTCGGTATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGT
GCTTTTAATCTAAATTGGGCTGATACTCAAGA

SEQ ID NO. 1404: SAG0471 FROM THE CJB110 GBS NONTYPEABLE STRAIN
TTGGTATCTTGACGCTTGAGGAGAAGTACAAGAAAAATGGGCAATTGAGACCAATACTTTAGAAAACGGAAGACATATCGTTTCTGATA
TCGTTGAATCTCTCAAACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGGTATCGGTATGGGGTCTCCAGGAGCTGTT
GATAGAACTAGTAAAAC

SEQ ID NO. 1405: SAG0471 FROM THE CJB110 GBS NONTYPEABLE STRAIN
CACCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATTGGTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGC
GTTGAGAAATACTTTGTCACATTTGCTTTCCCACAAGTTAAAAAGTCAACTA

SEQ ID NO. 1406: SAG0471 FROM THE 2603V/R GBS TYPE V STRAIN
GGGCAATTGAGACCAATACTTTAGAAAACGGAAGACATATCGTTTCTGATATCGTTGAATCTCTCAAACATCGTTTGAGCCTCTATGGA
TTAACAAAAGATGACTTTCTCGGTATCGGTATGGGTTCTCCAGGAGCTG

SEQ ID NO. 1407: SAG0471 FROM THE H36b GBS TYPE Ib STRAIN
GGCAATTGAGACCAATACTTTAGAAAACGGAAGACATATCGTTTCTGATATCGTTGAATCTCTCAAACATCGTTTGAGCCTCTATGGAT
TAACAAAAGATGACTTTCTCGGTATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGCTTTTAATCTA
AATTGGGCTGATACTCAAGAAGTAGGTTCAGTTATTGAAAAAGAAGTTGGAATTCCATTTTTTATTGATAACGATGCTAATGTTGCAGC
ACTTGGTGAACGCTGGGTAGGTGCTGGTGCCAATAATCCCGACGTTGTTTTCGTAACC

SEQ ID NO. 1408: SAG0471 FROM THE H36 GBS TYPE Ib STRAIN (REVERSE COMPLEMENT)
GAGACAGTTGCATCAGCGACAGGTGTTGTTAGAGTAGCACGTCAACTCGCAGAACAATATGAGGGTTCGTCTGCCATTAAAGCAGCGAT
TGACAACGGTGATACTGTTACAAGTAAAGATATTTTTATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTAT
CACGTTACCTTGGACTGGCAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATTGGTGGCGGTGTCTCAGCAGCAGGT
GAATTTTTACGTAGTCGCGTTGAGAAATACTTTGTCACATTTGCTTTCCCACA

SEQ ID NO. 1409: SAG0471 FROM THE M732 GBS TYPE III STRAIN
ACAAGAAAAATGGGCAATTGAGACCATACTTAGAAAACGGAAGACATATCGTTTCTGATATCGTTGAATCTCTCAAACATCGTTTGAGC
CTCTATGGATTAACAAAAGATGACTTTCTCGGTATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGC
TTTTAATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCGGTTATTGAAAAAGAAGTTGGAATTCCATTTTTTATTGATAACGATGCTA
ATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGCTGGTGCCAATAATCCCGATGTTGTTTTCGTAACCCTCGGAACAGGAGTAGGTGGA
GGTGTTATCGCAGATGGTAACCTCATCCATGGTGTTGCAAGAGCAGGTGGAGAAATTGGGCATATGATT

SEQ ID NO. 1410: SAG0471 FROM THE M732 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
CAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTTGTCACATTTGCTTTCCCACAAGTTAAAAAGTCAACTAAAATTAAG
ATTGCTGAACTAGGTAATGAT

SEQ ID NO. 1411: SAG0471 FROM THE M781 GBS TYPE III STRAIN
AGAAGTACAAGAAAATGGGCAATTGAGACCATACTTAGAAAACGGAAGACATATCGTTTCTGATATCGTTGAATCTCTCAAACATCGTT
TGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGGTATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACA
GGTGCTTTTAATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCGGTTATTGAAAAAGAAGTTGGAATTCCATTTTTTATTGATAACGA
TGCTAATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGCTGGTGCCAATAATCCCGATGTTGTTTTCGTAACCCTCGGAACAGGAGTA

Table 14: Comparative Sequences relating to SAG0471 (glucokinase)

SEQ ID NO. 1412: SAG0471 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
GATACTGTTACAAGTAAAGATATTTTTATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTATCACGTTACCT
TGGACTGGCAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATTGGTGGCGGTGTCTCAGCAGCAGGTGAATTTTTAC
GTAGTCGCGTTGAGAAATACTTTGTCACATTTGCTTTCCCACAAGTTAAAAA

SEQ ID NO. 1413: SAG0471 FROM THE 090 GBS TYPE Ia STRAIN
AAATTTGGTATCTTGACGCTTGAGGGAGAAGTACAAGAAAAATGGGCATTGAGACCATACTTAGAAAACGGAAGACATATCGTTTCTGA
TATCGTTGAATCTCTCAAACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGGTATCGGTATGGGTTCTCCAGGAGCTG
TTGATAGAACTAGTAAAACAGTAACAGGTGCTTTTAATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCAGTTATTGAAAAAGAAGTT
GGAATTCCATTTTTTATTGATAACGATGCTAATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGCTGGTGCCAATAATCCCGACGTTGT
TTTCGTAACCCTCGGAACAGGAGTAGGTGGAGG

SEQ ID NO. 1414: SAG0471 FROM THE 090 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
GTGATACTGTTACAAGTAAAGATATTTTTATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTATCACGTTAC
CTTGGACTGGCAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATTGGTGGCGGTGTCTCAGCAGCAGGTGAATTTTT
ACGTAGTCGCGTTGAGAAATACTTTATCACATTTGCTTTCCCACAAGTTAAAAAGTCAACTAAAATTAAGATTG

SEQ ID NO. 1415: SAG0471 FROM THE JM9130013 GBS TYPE VIII STRAIN (REVERSE COMPLEMENT)
GTTATCGCAGATGGTAACCTCATCCATGGTGTTGCAGGAGCAGGTGGAGAAATTGGGCATATGATTGTTGATCCAGAAAATGGATTTAC
GTGCACATGTGGTAACAAAGGCTGCCTTGAGACAGTTGCATCAGCGACAGGTGTTGTTAGAGTAGCACGTCAACTCGCAGAACAATATG
AGGGTTCGTCTGCCATTAAAGCAGCGATTGACCACGGTGATACTGTTACAAGTAAAGATATTTTTATAGCAGCAGAAGATGGGGATAAA
TTTGCTAATTCTGTTGTTGAACGTGTATCACGTTACCTTGGACTGGCAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGT
TATTGGTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTTGTCACATTTGCTTTCCCACAAGTTAAAA
AGTCAACTAA

SEQ ID NO. 1416: SAG0471 FROM THE JM9130013 GBS TYPE VIII STRAIN (REVERSE COMPLEMENT)
TGGTATCTTGACGCTTGAGGGAGAAGTACAAGAAAAATGGGCAATTGAGACCATACTTAGAAAACGGAAGACATATCGTTTCTGATATC
GTTGAATCTCTCAAACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGGTATCGGTATGGGTTCTCCAGGAGCTGTTGA
TAGAACTAGTAAAACAGTCACAGGTGCTTTTAATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCAGTTATTGAAAAAGAAGCTGGAA
TTCCATTTTTTATTG

SEQ ID NO. 1417: SAG0471 FROM THE 2603V/R TYPE V GBS STRAIN (REVERSE COMPLEMENT)
AGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATTGGTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCG
TTGAGAAATACTTTGTCACATTTGTTTTCCCACAAGGT

```
SEQ1401_        ------------------------------------------------------------
SEQ1402         ------------------------------------------------------------
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         ------------------------------------------------------------
SEQ1409         ------------------------------------------------------------
SEQ1410         ------------------------------------------------------------
SEQ1411         ------------------------------------------------------------
SEQ1412         ------------------------------------------------------------
SEQ1413         ------------------------------------------------------------
SEQ1414         ------------------------------------------------------------
SEQ1415         TTATCGCAGATGGTAACCTCATCCATGGTGTTGCAGGAGCAGGTGGAGAAATTGGGCAT
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------

SEQ1401_        ------------------------------------------------------------
SEQ1402         ------------------------------------------------------------
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         ---------------------------------------------------------GAG
SEQ1409         ------------------------------------------------------------
SEQ1410         ------------------------------------------------------------
SEQ1411         ------------------------------------------------------------
SEQ1412         ------------------------------------------------------------
SEQ1413         ------------------------------------------------------------
SEQ1414         ------------------------------------------------------------
SEQ1415         TGATTGTTGATCCAGAAAATGGATTTACGTGCACATGTGGTAACAAAGGCTGCCTTGAG
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------
```

Table 14: Comparative Sequences relating to SAG0471 (glucokinase)

```
SEQ1401_        ------------------------------------------------------------
SEQ1402         ------------------------------------------------------------
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         CAGTTGCATCAGCGACAGGTGTTGTTAGAGTAGCACGTCAACTCGCAGAACAATATGAG
SEQ1409         ------------------------------------------------------------
SEQ1410         ------------------------------------------------------------
SEQ1411         ------------------------------------------------------------
SEQ1412         ------------------------------------------------------------
SEQ1413         ------------------------------------------------------------
SEQ1414         ------------------------------------------------------------
SEQ1415         CAGTTGCATCAGCGACAGGTGTTGTTAGAGTAGCACGTCAACTCGCAGAACAATATGAG
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------

SEQ1401_        ------------------------------------------------------------
SEQ1402         ------------------------------------------------------------
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         GTTCGTCTGCCATTAAAGCAGCGATTGACAACGGTGATACTGTTACAAGTAAAGATATT
SEQ1409         ------------------------------------------------------------
SEQ1410         ------------------------------------------------------------
SEQ1411         ------------------------------------------------------------
SEQ1412         ------------------------------------GATACTGTTACAAGTAAAGATATT
SEQ1413         ------------------------------------------------------------
SEQ1414         ----------------------------------GTGATACTGTTACAAGTAAAGATATT
SEQ1415         GTTCGTCTGCCATTAAAGCAGCGATTGACCACGGTGATACTGTTACAAGTAAAGATATT
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------

SEQ1401_        ----------------------TTAAATTTGGTATCTTGACGCTTGAGGGAGAAGTACAA
SEQ1402         ------------------------------------------------------------
SEQ1403         --------------------------------------------------------ACAA
SEQ1404         ----------------------------TTGGTATCTTGACGCTTGAGG-AGAAGTACAA
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         TTATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTATCACGT
SEQ1409         --------------------------------------------------------ACAA
SEQ1410         ------------------------------------------------------------
SEQ1411         --------------------------------------------------AGAAGTACAA
SEQ1412         TTATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTATCACGT
SEQ1413         ------------------------AAATTTGGTATCTTGACGCTTGAGGGAGAAGTACAA
SEQ1414         TTATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTATCACGT
SEQ1415         TTATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTATCACGT
SEQ1416         -----------------------------TGGTATCTTGACGCTTGAGGGAGAAGTACAA
SEQ1417         ------------------------------------------------------------
```

Table 14: Comparative Sequences relating to SAG0471 (glucokinase)

```
SEQ1401_        AAAAATGGGCAATTGAGACCAATACTTTAGAAAACGGAAGACATATCGTTTCTGATATC
SEQ1402         -----------------------------------------------CGTTTCTGATATC
SEQ1403         AAAAATGGGCAATTGAGACCAATACTTTAGAAAACGGAAGACATATCGTTTCTGATATC
SEQ1404         AAAAATGGGCAATTGAGACCAATACTTTAGAAAACGGAAGACATATCGTTTCTGATATC
SEQ1405         ------------CACCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATT
SEQ1406         ------GGGCAATTGAGACCAATACTTTAGAAAACGGAAGACATATCGTTTCTGATATC
SEQ1407         -------GGCAATTGAGACCAATACTTTAGAAAACGGAAGACATATCGTTTCTGATATC
SEQ1408         ACCTTGGACTGGCAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATT
SEQ1409         AAAAATGGGCAATTGAGACCA-TACTT-AGAAAACGGAAGACATATCGTTTCTGATATC
SEQ1410         -----------------------------------------------------------
SEQ1411         AAAA-TGGGCAATTGAGACCA-TACTT-AGAAAACGGAAGACATATCGTTTCTGATATC
SEQ1412         ACCTTGGACTGGCAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATT
SEQ1413         AAAAATGGGCA-TTGAGACCA-TACTT-AGAAAACGGAAGACATATCGTTTCTGATATC
SEQ1414         ACCTTGGACTGGCAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATT
SEQ1415         ACCTTGGACTGGCAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATT
SEQ1416         AAAAATGGGCAATTGAGACCA-TACTT-AGAAAACGGAAGACATATCGTTTCTGATATC
SEQ1417         -------------AGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATT

SEQ1401_        TTGAATCTCTCA-AACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGG
SEQ1402         TTGAATCTCTCA-AACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGG
SEQ1403         TTGAATCTCTCA-AACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGG
SEQ1404         TTGAATCTCTCA-AACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGG
SEQ1405         GTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTT---
SEQ1406         TTGAATCTCTCA-AACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGG
SEQ1407         TTGAATCTCTCA-AACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGG
SEQ1408         GTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTT---
SEQ1409         TTGAATCTCTCA-AACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGG
SEQ1410         ------------CAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTT---
SEQ1411         TTGAATCTCTCA-AACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGG
SEQ1412         GTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTT---
SEQ1413         TTGAATCTCTCA-AACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGG
SEQ1414         GTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTT---
SEQ1415         GTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTT---
SEQ1416         TTGAATCTCTCA-AACATCGTTTGAGCCTCTATGGATTAACAAAAGATGACTTTCTCGG
SEQ1417         GTGGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTT---

SEQ1401_        ATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGCTTT
SEQ1402         ATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGCTTT
SEQ1403         ATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGCTTT
SEQ1404         ATCGGTATGGGGTCTCCAGGAGCTGTTGATAGAACTAGTAAAAC---------------
SEQ1405         GTCACATTTGCTTTCCCACAAGTTAAAAAGTCAACTA----------------------
SEQ1406         ATCGGTATGGGTTCTCCAGGAGCTG----------------------------------
SEQ1407         ATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGCTTT
SEQ1408         GTCACATTTGCTTTCCCACA---------------------------------------
SEQ1409         ATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGCTTT
SEQ1410         GTCACATTTGCTTTCCCACAAGTTAAAAAGTCAACTAAAATTAAGATTGCTGAACTAGG
SEQ1411         ATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGCTTT
SEQ1412         GTCACATTTGCTTTCCCACAAGTTAAAAA------------------------------
SEQ1413         ATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTAACAGGTGCTTT
SEQ1414         ATCACATTTGCTTTCCCACAAGTTAAAAAGTCAACTAAAATTAAGATTG----------
SEQ1415         GTCACATTTGCTTTCCCACAAGTTAAAAAGTCAACTAA---------------------
SEQ1416         ATCGGTATGGGTTCTCCAGGAGCTGTTGATAGAACTAGTAAAACAGTCACAGGTGCTTT
SEQ1417         GTCACATTTGTTTTCCCACAAGGT-----------------------------------
```

Table 14: Comparative Sequences relating to SAG0471 (glucokinase)

```
SEQ1401_        AATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCAGTTATTGAAAAAGAAGTTGGAAT
SEQ1402         AATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCGGTTATTGAAAAAGAAGTTGGAAT
SEQ1403         AATCTAAATTGGGCTGATACTCAAGA----------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         AATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCAGTTATTGAAAAAGAAGTTGGAAT
SEQ1408         ------------------------------------------------------------
SEQ1409         AATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCGGTTATTGAAAAAGAAGTTGGAAT
SEQ1410         AATGAT------------------------------------------------------
SEQ1411         AATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCGGTTATTGAAAAAGAAGTTGGAAT
SEQ1412         ------------------------------------------------------------
SEQ1413         AATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCAGTTATTGAAAAAGAAGTTGGAAT
SEQ1414         ------------------------------------------------------------
SEQ1415         ------------------------------------------------------------
SEQ1416         AATCTAAATTGGGCTGATACTCAAGAAGTAGGTTCAGTTATTGAAAAAGAAGCTGGAAT
SEQ1417         ------------------------------------------------------------

SEQ1401_        CCATTTTTTTATTGATAACGATGCTAATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGC
SEQ1402         CCATTTTTTTATTGATAACGATGCTAATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGC
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         CCATTTTTTTATTGATAACGATGCTAATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGC
SEQ1408         ------------------------------------------------------------
SEQ1409         CCATTTTTTTATTGATAACGATGCTAATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGC
SEQ1410         ------------------------------------------------------------
SEQ1411         CCATTTTTTTATTGATAACGATGCTAATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGC
SEQ1412         ------------------------------------------------------------
SEQ1413         CCATTTTTTTATTGATAACGATGCTAATGTTGCAGCACTTGGTGAACGCTGGGTAGGTGC
SEQ1414         ------------------------------------------------------------
SEQ1415         ------------------------------------------------------------
SEQ1416         CCATTTTTTTATTG----------------------------------------------
SEQ1417         ------------------------------------------------------------

SEQ1401_        GGTGCCAATAATCCCGACGTTGTTTTCGTAACCCTCGGAACAGGAGTAGGTGGAGGTGT
SEQ1402         GGTGCCAATAATCCCGATGTTGTTTTCGTAACCCTCGGAACAGGAGTAGGTGGAGGTGT
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         GGTGCCAATAATCCCGACGTTGTTTTCGTAACC---------------------------
SEQ1408         ------------------------------------------------------------
SEQ1409         GGTGCCAATAATCCCGATGTTGTTTTCGTAACCCTCGGAACAGGAGTAGGTGGAGGTGT
SEQ1410         ------------------------------------------------------------
SEQ1411         GGTGCCAATAATCCCGATGTTGTTTTCGTAACCCTCGGAACAGGAGTA-----------
SEQ1412         ------------------------------------------------------------
SEQ1413         GGTGCCAATAATCCCGACGTTGTTTTCGTAACCCTCGGAACAGGAGTAGGTGGAGG---
SEQ1414         ------------------------------------------------------------
SEQ1415         ------------------------------------------------------------
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------
```

Table 14: Comparative Sequences relating to SAG0471 (glucokinase)

```
SEQ1401_        ATCGCAGATGGTAACCTCATCCATGGTGTTGCAGGAGCAGGTGGAGAAATTGGGCATAT
SEQ1402         ATCGCAGATGGTAACCTCATCCATGGTGTTGCAGGAGCAGGTGGAGAAATTGGGCATAT
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         ------------------------------------------------------------
SEQ1409         ATCGCAGATGGTAACCTCATCCATGGTGTTGCAAGAGCAGGTGGAGAAATTGGGCATAT
SEQ1410         ------------------------------------------------------------
SEQ1411         ------------------------------------------------------------
SEQ1412         ------------------------------------------------------------
SEQ1413         ------------------------------------------------------------
SEQ1414         ------------------------------------------------------------
SEQ1415         ------------------------------------------------------------
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------

SEQ1401_        ATTGTTGATCCAGAAAATGGATTTACGTGCACATGTGGTAACAAAGGCTGCCTTGAGAC
SEQ1402         ATTGTTGATCCAGAKAATGGATTTACGTGCACATGTGGTAACAAAGGCTGTCTTGAGAC
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         ------------------------------------------------------------
SEQ1409         ATT---------------------------------------------------------
SEQ1410         ------------------------------------------------------------
SEQ1411         ------------------------------------------------------------
SEQ1412         ------------------------------------------------------------
SEQ1413         ------------------------------------------------------------
SEQ1414         ------------------------------------------------------------
SEQ1415         ------------------------------------------------------------
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------

SEQ1401_        GTTGCATCAGCGACAGGTGTTGTTAGAGTAGCACGTCAACTCGCAGAACAATATGAGGG
SEQ1402         GTTGCATCAGCGACAGGTGTTGTTAGAGTAGCACGTCAACTCGCAGAACAATATGAAGG
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         ------------------------------------------------------------
SEQ1409         ------------------------------------------------------------
SEQ1410         ------------------------------------------------------------
SEQ1411         ------------------------------------------------------------
SEQ1412         ------------------------------------------------------------
SEQ1413         ------------------------------------------------------------
SEQ1414         ------------------------------------------------------------
SEQ1415         ------------------------------------------------------------
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------
```

Table 14: Comparative Sequences relating to SAG0471 (glucokinase)

```
SEQ1401_        TCGTCTGCCATTAAAGCAGCGATTGACACCGGTGATACTGTTACAAGTAAAGATATTTT
SEQ1402         TCGTCTGCCATTAAAGCAGCGATTGACAACGGTGATACTGTTACAAGTAAAGATATTTT
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         ------------------------------------------------------------
SEQ1409         ------------------------------------------------------------
SEQ1410         ------------------------------------------------------------
SEQ1411         ------------------------------------------------------------
SEQ1412         ------------------------------------------------------------
SEQ1413         ------------------------------------------------------------
SEQ1414         ------------------------------------------------------------
SEQ1415         ------------------------------------------------------------
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------

SEQ1401_        ATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTATCACGTTA
SEQ1402         ATAGCAGCAGAAGATGGGGATAAATTTGCTAATTCTGTTGTTGAACGTGTATCACGTTA
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         ------------------------------------------------------------
SEQ1409         ------------------------------------------------------------
SEQ1410         ------------------------------------------------------------
SEQ1411         ------------------------------------------------------------
SEQ1412         ------------------------------------------------------------
SEQ1413         ------------------------------------------------------------
SEQ1414         ------------------------------------------------------------
SEQ1415         ------------------------------------------------------------
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------

SEQ1401_        CTTGGACTGGCAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATTGG
SEQ1402         CTTGGACTGGCAGCAGCTAATATTTCAAATATTTTAAACCCTGATTCTGTGGTTATTGG
SEQ1403         ------------------------------------------------------------
SEQ1404         ------------------------------------------------------------
SEQ1405         ------------------------------------------------------------
SEQ1406         ------------------------------------------------------------
SEQ1407         ------------------------------------------------------------
SEQ1408         ------------------------------------------------------------
SEQ1409         ------------------------------------------------------------
SEQ1410         ------------------------------------------------------------
SEQ1411         ------------------------------------------------------------
SEQ1412         ------------------------------------------------------------
SEQ1413         ------------------------------------------------------------
SEQ1414         ------------------------------------------------------------
SEQ1415         ------------------------------------------------------------
SEQ1416         ------------------------------------------------------------
SEQ1417         ------------------------------------------------------------
```

Table 14: Comparative Sequences relating to SAG0471 (glucokinase)

```
SEQ1401_        GGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTTGTCAC
SEQ1402         GGCGGTGTCTCAGCAGCAGGTGAATTTTTACGTAGTCGCGTTGAGAAATACTTTGTCAC
SEQ1403         -----------------------------------------------------------
SEQ1404         -----------------------------------------------------------
SEQ1405         -----------------------------------------------------------
SEQ1406         -----------------------------------------------------------
SEQ1407         -----------------------------------------------------------
SEQ1408         -----------------------------------------------------------
SEQ1409         -----------------------------------------------------------
SEQ1410         -----------------------------------------------------------
SEQ1411         -----------------------------------------------------------
SEQ1412         -----------------------------------------------------------
SEQ1413         -----------------------------------------------------------
SEQ1414         -----------------------------------------------------------
SEQ1415         -----------------------------------------------------------
SEQ1416         -----------------------------------------------------------
SEQ1417         -----------------------------------------------------------

SEQ1401_        TTTGCTTTCCCACAAGTTAAAAAGTCAACTAAAATTAAGAT
SEQ1402         TTTG-------------------------------------
SEQ1403         -----------------------------------------
SEQ1404         -----------------------------------------
SEQ1405         -----------------------------------------
SEQ1406         -----------------------------------------
SEQ1407         -----------------------------------------
SEQ1408         -----------------------------------------
SEQ1409         -----------------------------------------
SEQ1410         -----------------------------------------
SEQ1411         -----------------------------------------
SEQ1412         -----------------------------------------
SEQ1413         -----------------------------------------
SEQ1414         -----------------------------------------
SEQ1415         -----------------------------------------
SEQ1416         -----------------------------------------
SEQ1417         -----------------------------------------
```

Table 15: Comparative Sequences relating to SAG0492

SEQ ID NO. 1501: SAG0492 FROM THE 1169NT1 GBS NONTYPEABLE STRAIN
TGACTTGGATATTCATCAAGGAGAAGTGGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAAGAACAATGAATCTCT
TGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGAATTGATATAACAGACAAAAAAAATGATATTTTTAAAATGCGCGAAAAAATG
GGCATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTATTAAGACAAAGGGACTTTC
TAAGCTTGATGCTCAGACAAAAGCATACGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCTAGCTTATCTG
GAGGACAACAACAACGGATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAACCTACTTCAGCTCTTGAT
CCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAGCTAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTT
TGCACGTGAAGTAGCGGATCGTGTCATTTTTATGGATGCAGGCATTATTGTGAGCAAGGGACCCCTAAGGAAGTAT

SEQ ID NO. 1502: SAG0492 FROM THE 18RS21 GBS TYPE II STRAIN
TTGGGAAAAAATGAGGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTAGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAG
TCAACATTTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTGATATAACAGACAAAAAGAA
TGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTT
TATCACCTATTAAGACAAAGGGGCTTTCTAATCTTGATGCTCAGACAAAAGCATATGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAG
GCTAATACTTATCCAGCTAGCTTATCTGGAGGACAACAACAACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTCATGTCCTTCT
TTTTGATGAACCTACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAGCTAAATCTGGTATGACGA
TGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAGCGGATCGTGTCATTTTTATGGACGCAGAAATTAT

SEQ ID NO. 1503: SAG0492 FROM THE 2603V/R GBS TYPE V STRAIN (REVERSE COMPLEMENT)
AAAAATGAGGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTAGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAAC
ATTTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTGATATAACAGACAAAAAGAATGATA
TTTTTAAAATGCGCGAAAAAATGGGCATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCA
CCTATTAAGACAAAGGGGCTTTCTAATCTTGATGCTCAGACAAAAGCATATGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAA
TACTTATCCAGCTAGCTTATCTGGAGGACAACAACAACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTG
ATGAACCTACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAGCTAAATCTGGTATGACGATGGTT
ATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAGCGGATCGTGTCATTTTTATGGATGCAGGAATTATTGTTGAGCAAGGGGCCC

SEQ ID NO. 1504: SAG0492 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
GAGGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTGGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTT
AAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTGATATAACAGACAAAAAGAATGATATTTTTA
AAATGCGCGAAAAAATGGGCATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTATT
AAGACAAAGGGACTTTCTAAGCTTGATGCTCAGACAAAAGCATACGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAATGCTTA
TCCAGCAAGCTTATCTGGAGGACAACAACAACGGATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAAC
CTACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAGCTAAATCTGGTATGACGATGGTTATTGTC
ACTCATGAAATGGGTTTTGCACGTGAAGTAGCGGATCGTGTCATTTTTATGGATGCAGGGATTATTGTTGAGCAAGGGACCCCTAAGAA
AGTAT

SEQ ID NO. 1505: SAG0492 FROM THE 090 GBS TYPE Ia STRAIN
TGGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTG
ACTTTTGAAGGGATTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTTTTCAACAGTTCAATCT
ATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTATTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAGACAAAAGCAT
ACGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCTAGCTTATCTGGAGGGCAACAACAACGAATTGCTATT
GCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAACCTACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGAC
TGTTATGCAAGATTTAGCTAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAGCGGATCGTGTCA
TTTTTATGGATGCAGGCATTATTGTTgAsCAAGGGACCCCTAAGGAAGTA

SEQ ID NO. 1506: SAG0492 FROM THE A909 GBS TYPE Ia STRAIN
CAATACAAGGACTTCATAAAAGTTTTGGGAAAAATGAGGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTAGTGGTTATT
ATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGG
GATTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTTTTCAACAGTTCAATCTATTTCCCAATA
TGACTGTACTAGAAAATATTACTTTATCACCTATTAAGACAAAGGGGCTTTCTAAGCTTGATGCTCAGACAAAAGCATATGAGCTACTT
GAAAAAGTTGGACTCAAAGAGAAGGCTAATACTTATCCAGCTAGCTTATCTGGAGGACAACAACAACGAATTGCTATTGCAAGAGGTCT
TGCAATGAATCCTGATGTCCTTCTTTTTGATGAACCTACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAG
ATTTAGCTAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAGCGGATCGTGTCATTTTTATGGAT
GCAGGAATTATTGTgAGCAAGGGGCCCCTAAGGAAGTATTTGAGCAGACAAAAGAAATCCGCACAAGAGATTTCTT

SEQ ID NO. 1507: SAG0492 FROM THE CJB110 GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)
GACTTGGATATTCATCAAGGAGAAGTGGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAAGAACAATGAATCTCTT
GGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGG
GCATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTATTAAGACAAAGGGACTTTCT
AAGCTTGATGCTCAGACAAAAGCATACGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCTAGCTTATCTGG
AGGACAACAACAACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAACCTACTTCAGCTCTTGATC
CTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAGCTAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTT
GCACGTGAAGTAGCGGATCGTGTCTTTTTATGGATGCGGGAATTATTGTGAGCAAGGGACC

Table 15: Comparative Sequences relating to SAG0492

SEQ ID NO. 1508: SAG0492 FROM THE H36b GBS TYPE Ib STRAIN
ATGAGGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTAGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTT
TTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTGATATAACAGACAAAAAGAATGATATTTT
TAAAATGCGCGAAAAAATGGGCATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
TTAAGACAAAGGGGCTTTCTAAGCTTGATGCTCAGACAAAAGCATATGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAATACT
TATCCAGCTAGCTTATCTGGAGGACAACAACAACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGA
ACCTACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAGCTAAATCTGGTATGACGATGGTTATTG
TCACTCATGAAATGGGTTTTGCACGTGAAGTAGCGGATCGTGTCATTTTTATGGATGCASGAATTATTGTTGAGCAAGGGGCCCCTAAG
GAAGTAT

SEQ ID NO. 1509: SAG0492 FROM THE JM9130013 GBS TYPE VIII STRAIN (REVERSE COMPLEMENT)
GGTTTTAAAAGGCATTGACTTGGATATTCATCAAGGAGAAGTAGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAA
GAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTGATATAACAGACAAAAAGAATGATATTTTTAAA
ATGCGCGAAAAAATGGGCATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTATTAA
GACAAAGGGGCTTTCTAAGCTTGATGCTCAGACAAAAGCATATGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAATACTTATC
CAGCTAGCTTATCTGGAGGACAACAACAACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAACCT
ACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAGCTAAATCTGGTATGACGATGGTTATTGTCAC
TCATGAAATGGGTTTTGCACGTGAAGTAGCGGATCGTGTCATTTTTATGGATGCAGGAATTATTGTTGAGCAAGGGGCCCCTAAGGAAG
TATTTAGCAAAACAAAAGAAAT

SEQ ID NO. 1510: SAG0492 FROM THE M732 GBS TYPE III STRAIN
GGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGA
CTTTTGAAGGGATTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTTTTCAACAGTTCAATCTA
TTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTATTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAGACAAAAGCATA
CGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCAAGCTTATCTGG

SEQ ID NO. 1511: SAG0492 FROM THE COH1 GBS TYPE Ia STRAIN
ATTGACTTGGATATTCATCAAGGAGAAGTGGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACATTTTTAAGAACAATGAATCT
CTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGATTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAA
TGGGCATGGTTTTTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTATTAAGACAAAGGGACTT
TCTAAGCTTGATGCTCAGACAAAAGCATACGAGCTACTTGAAAAAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCAAGCTTATC
TGG

Table 15: Comparative Sequences relating to SAG0492

```
SEQ1501    ---------------------------------------------------TGACTTGG
SEQ1502    -----------------------TTGGGAAAAATGAGGTTTTAAAAGGCATTGACTTGG
SEQ1503    ----------------------------AAAAATGAGGTTTTAAAAGGCATTGACTTGG
SEQ1504    ----------------------------------GAGGTTTTAAAAGGCATTGACTTGG
SEQ1505    -----------------------------------------------------------
SEQ1506    AATACAAGGACTTCATAAAAGTTTTGGGAAAAATGAGGTTTTAAAAGGCATTGACTTGG
SEQ1507    ----------------------------------------------------GACTTGG
SEQ1508    --------------------------------ATGAGGTTTTAAAAGGCATTGACTTGG
SEQ1509    ------------------------------------GGTTTTAAAAGGCATTGACTTGG
SEQ1510    -----------------------------------------------------------
SEQ1511    -------------------------------------------------ATTGACTTGG

SEQ1501    TATTCATCAAGGAGAAGTGGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT
SEQ1502    TATTCATCAAGGAGAAGTAGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT
SEQ1503    TATTCATCAAGGAGAAGTAGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT
SEQ1504    TATTCATCAAGGAGAAGTGGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT
SEQ1505    -----------------TGGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT
SEQ1506    TATTCATCAAGGAGAAGTAGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT
SEQ1507    TATTCATCAAGGAGAAGTGGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT
SEQ1508    TATTCATCAAGGAGAAGTAGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT
SEQ1509    TATTCATCAAGGAGAAGTAGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT
SEQ1510    ------------------GGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT
SEQ1511    TATTCATCAAGGAGAAGTGGTGGTTATTATTGGCCCTTCTGGCTCTGGTAAGTCAACAT

SEQ1501    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGAA
SEQ1502    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGA
SEQ1503    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGA
SEQ1504    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGA
SEQ1505    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGA
SEQ1506    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGA
SEQ1507    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGA
SEQ1508    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGA
SEQ1509    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGA
SEQ1510    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGA
SEQ1511    TTTTAAGAACAATGAATCTCTTGGAAGTACCAACAAAGGGAACAGTGACTTTTGAAGGGA

SEQ1501    TTGATATAACAGACAAAAAAAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT
SEQ1502    TTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT
SEQ1503    TTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT
SEQ1504    TTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT
SEQ1505    TTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT
SEQ1506    TTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT
SEQ1507    TTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT
SEQ1508    TTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT
SEQ1509    TTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT
SEQ1510    TTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT
SEQ1511    TTGATATAACAGACAAAAAGAATGATATTTTTAAAATGCGCGAAAAAATGGGCATGGTTT

SEQ1501    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
SEQ1502    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
SEQ1503    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
SEQ1504    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
SEQ1505    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
SEQ1506    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
SEQ1507    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
SEQ1508    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
SEQ1509    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
SEQ1510    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
SEQ1511    TTCAACAGTTCAATCTATTTCCCAATATGACTGTACTAGAAAATATTACTTTATCACCTA
```

Table 15: Comparative Sequences relating to SAG0492

```
SEQ1501     TTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAGACAAAAGCATACGAGCTACTTGAAA
SEQ1502     TTAAGACAAAGGGGCTTTCTAATCTTGATGCTCAGACAAAAGCATATGAGCTACTTGAAA
SEQ1503     TTAAGACAAAGGGGCTTTCTAATCTTGATGCTCAGACAAAAGCATATGAGCTACTTGAAA
SEQ1504     TTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAGACAAAAGCATACGAGCTACTTGAAA
SEQ1505     TTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAGACAAAAGCATACGAGCTACTTGAAA
SEQ1506     TTAAGACAAAGGGGCTTTCTAAGCTTGATGCTCAGACAAAAGCATATGAGCTACTTGAAA
SEQ1507     TTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAGACAAAAGCATACGAGCTACTTGAAA
SEQ1508     TTAAGACAAAGGGGCTTTCTAAGCTTGATGCTCAGACAAAAGCATATGAGCTACTTGAAA
SEQ1509     TTAAGACAAAGGGGCTTTCTAAGCTTGATGCTCAGACAAAAGCATATGAGCTACTTGAAA
SEQ1510     TTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAGACAAAAGCATACGAGCTACTTGAAA
SEQ1511     TTAAGACAAAGGGACTTTCTAAGCTTGATGCTCAGACAAAAGCATACGAGCTACTTGAAA

SEQ1501     AAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCTAGCTTATCTGGAGGACAACAAC
SEQ1502     AAGTTGGACTCAAAGAGAAGGCTAATACTTATCCAGCTAGCTTATCTGGAGGACAACAAC
SEQ1503     AAGTTGGACTCAAAGAGAAGGCTAATACTTATCCAGCTAGCTTATCTGGAGGACAACAAC
SEQ1504     AAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCAAGCTTATCTGGAGGACAACAAC
SEQ1505     AAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCTAGCTTATCTGGAGGGCAACAAC
SEQ1506     AAGTTGGACTCAAAGAGAAGGCTAATACTTATCCAGCTAGCTTATCTGGAGGACAACAAC
SEQ1507     AAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCTAGCTTATCTGGAGGACAACAAC
SEQ1508     AAGTTGGACTCAAAGAGAAGGCTAATACTTATCCAGCTAGCTTATCTGGAGGACAACAAC
SEQ1509     AAGTTGGACTCAAAGAGAAGGCTAATACTTATCCAGCTAGCTTATCTGGAGGACAACAAC
SEQ1510     AAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCAAGCTTATCTGG-----------
SEQ1511     AAGTTGGACTCAAAGAGAAGGCTAATGCTTATCCAGCAAGCTTATCTGGTABCMARATVS

SEQ1501     ACGGATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAAC
SEQ1502     ACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTCATGTCCTTCTTTTTGATGAAC
SEQ1503     ACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAAC
SEQ1504     ACGGATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAAC
SEQ1505     ACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAAC
SEQ1506     ACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAAC
SEQ1507     ACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAAC
SEQ1508     ACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAAC
SEQ1509     ACGAATTGCTATTGCAAGAGGTCTTGCAATGAATCCTGATGTCCTTCTTTTTGATGAAC
SEQ1510     -----------------------------------------------------------
SEQ1511     NCSRATNGTSAG-----------------------------------------------

SEQ1501     TACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAG
SEQ1502     TACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAG
SEQ1503     TACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAG
SEQ1504     TACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAG
SEQ1505     TACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAG
SEQ1506     TACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAG
SEQ1507     TACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAG
SEQ1508     TACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAG
SEQ1509     TACTTCAGCTCTTGATCCTGAAATGGTAGGTGAAGTCTTGACTGTTATGCAAGATTTAG
SEQ1510     -----------------------------------------------------------
SEQ1511     -----------------------------------------------------------

SEQ1501     TAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAG
SEQ1502     TAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAG
SEQ1503     TAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAG
SEQ1504     TAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAG
SEQ1505     TAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAG
SEQ1506     TAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAG
SEQ1507     TAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAG
SEQ1508     TAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAG
SEQ1509     TAAATCTGGTATGACGATGGTTATTGTCACTCATGAAATGGGTTTTGCACGTGAAGTAG
SEQ1510     -----------------------------------------------------------
SEQ1511     -----------------------------------------------------------
```

Table 15: Comparative Sequences relating to SAG0492

```
SEQ1501        GGATCGTGTCATTTTTATGGATGCAGGCATTATTGT-GAGCAAGGGACCCCTAAGGAAG
SEQ1502        GGATCGTGTCATTTTTATGGACGCAGAAATTAT--------------------------
SEQ1503        GGATCGTGTCATTTTTATGGATGCAGGAATTATTGTTGAGCAAGGGGCCC---------
SEQ1504        GGATCGTGTCATTTTTATGGATGCAGGGATTATTGTTGAGCAAGGGACCCCTAAGAAAG
SEQ1505        GGATCGTGTCATTTTTATGGATGCAGGCATTATTGTTGASCAAGGGACCCCTAAGGAAG
SEQ1506        GGATCGTGTCATTTTTATGGATGCAGGAATTATTGTGAGCAAGGGGCCCCTAAGGAAGT
SEQ1507        GGATCGTGTC-TTTTTATGGATGCGGGAATTATTGT-GAGCAAGGGACC----------
SEQ1508        GGATCGTGTCATTTTTATGGATGCASGAATTATTGTTGAGCAAGGGGCCCCTAAGGAAG
SEQ1509        GGATCGTGTCATTTTTATGGATGCAGGAATTATTGTTGAGCAAGGGGCCCCTAAGGAAG
SEQ1510        -----------------------------------------------------------
SEQ1511        -----------------------------------------------------------

SEQ1501        AT------------------------------------
SEQ1502        --------------------------------------
SEQ1503        --------------------------------------
SEQ1504        AT------------------------------------
SEQ1505        A-------------------------------------
SEQ1506        TTTGAGCAGACAAAAGAAATCCGCACAAGAGATTTCTT
SEQ1507        --------------------------------------
SEQ1508        AT------------------------------------
SEQ1509        ATTTAGCAAAACAAAAGAAAT-----------------
SEQ1510        --------------------------------------
SEQ1511        --------------------------------------
```

Table 16: Comparative Sequences relating to SAG0767 (D-alanine – D-alanine ligase)

SEQ ID NO. 1601: SAG0767 FROM THE M781 GBS TYPE III STRAIN

TGGTCGCTCTGTCGGAACGTGAAGTATCTGTACTGTCTGCAGAAAGCGTCATGCGTGCTATTAATTATGATAAATTTTTT
GTTAAAACTTATTTTATCACGCAAGTAGGTCAATTTATTAAAACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAA
GTTAATGACAAACCAAACTGTTGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATGATGATAATGCAATTGTTTTCC
CCGTTTTACATGGACCAATGGGGGAAGATGGTTCTATCCAAGGATTTTTAGAAGTTTTAAGGATGCCTTATGTTGGGACT
AATATTCTATCTTCAAGCGTGGCTATGGATAAAATTACAACAAAACAAGTCCTTGCAACAGTAGGTGTACCTCAGGTTGC
ATATCAAACTTATTTTGAGGGTGATGATTTGGAACATGCGATTAAACTCTCTTTAGAAACTTTAAGTTTCCCAATTTTTG
TAAAACCGGCTAATATGGGGTCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCACTTCGCTCTGCAATTGACTTA
GCTCTCAAGTATGATAGCCGTATTTTGATTGAACAAGGCGTGACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAA
TGATGTTAAGACAACTTTTCCTGGCGAAGTTGTTAAAGACGTCGATTTCTATGACTATGACGCCAAATATATTGATAATA
AAATTACTATGGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCGTCAATATGCAAGTAAAGCTTTTAAA
GCAATCGGGGCTTGTGGTTTATCACGCTGTGATTTCTTTTTGACGAAAGATGGACAAATCTTCTTAAACGAACTGAATAC
AATGCCCGGTTTTACTCAGTGGTCAATGTATCCTCTGCTTTGGGAAAATATGGGGCTAACTTATAGTGATTTGATTG

SEQ ID NO. 1602: SAG0767 FROM THE 090 GBS TYPE Ia STRAIN

AAACCGGGCATTGTATTCAGTTCGTTTAAGAAGACTTGTCCATCTTTCGTCAAAAAGAAATCACAGCGTGATAAACCACA
AGCCCCGATTGCTTTAAAAGCTTTACTTGCATATTGACGCATTGCTTCCATAGTTGCTTCATCAACTTTAGCTGGAATAT
CCATAGTAATTTTATTATCAATATATTTGGCGTCATAGTCATAGAAATCGACGTCTTTAACGACTTCGCCAGGAAAAGTT
GTCTTAACATCATTATTGCCTAAAATACCTACTTCAATTTCACGAGCTGTCACGCCTTGTTCAATCAAAATACGGCTATC
ATACTTGAGAGCTAAGTCAATksCAGAGCGAAGTGAGGATTCATCTGTCGCTTTTGAAATACCTACTGATGACCCCATAT
TAGCCGGTTTTACAAAAATTGGGAAACTTAAAGTTTCTAAAGAGAGTTTAATCGCATGTTCCAAATCATCACCCTCAAAA
TAAGTTTGATATGCAACCTGAGGTACACCTACTGTTGCAAGGACTTGTTTTGTTGTAATTTTATCCATAGCCACGCTTGA
AGATAGAATATTAGTCCCAACATAAGGCATCCTTAAAACTTCTAAAAATCCTTGGATAGAACCATCTTCCCCCATTGGTC
CATGTAAAACGGGGAAAACAATTGCATTATCATCATAGATATCACTTGGACGAACCATTTTGTCTAAATCAACAGTTTGG
TTTGTCATTAACTTTTCATCTGAAGATGGCATTTCATCAAATTCTTGTGTTTTAATAAATTGACCTACTTGCGTG

SEQ ID NO. 1603: SAG0767 FROM THE COH1 TYPE Ia STRAIN

TCGCTCTGCGGAACGTGAAGTATCTGTACTGTCTGCAGAAAGCGTCATGCGTGCTATTAATTATGATAAATTTTTTGTTA
AAACTTATTTTATCACGCAAGTAGGTCAATTTATTAAAACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTA
ATGACAAACCAAACTGTTGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATGATGATAATGCAATTGTTTTCCCCGT
TTTACATGGACCAATGGGGGAAGATGGTTCTATCCAAGGATTTTTAGAAGTTTTAAGGATGCCTTATGTTGGGACTAATA
TTCTATCTTCAAGCGTGGCTAT

SEQ ID NO. 1604: SAG0767 FROM THE CJB110 GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)

CGTCGATTTCTATGACTATGACGCCAAATATATTGATAATAAAATTACTATGGATATTCCAGCTAAAGTTGATGAAGCAA
CTATGGAAGCAATGCGTCAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTTCTTT
TTGACGAAAGATGGACAAATCTTCTTAAACGAACTGAATACAATGCCC

SEQ ID NO. 1605: SAG0767 FROM THE CJB110 GBS NONTYPEABLE STRAIN

AACGTGAAGTATCTGTACTGCTCTGCAGAAAAGCGTCATGCGTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATT
TTATCACGCAAGTAGGTCAATTTATTAAAACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAA

SEQ ID NO. 1606: SAG0767 FROM THE 1169NT1 GBS TYPE V STRAIN (REVERSE COMPLEMENT)

CTAATATGGGGTCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCACTTCGCTCTGCAATTGACTTAGCTCTCAAG
TATGATAGCCGTATTTTGATTGAACAAGGCGTGACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAA
GACAACTTTTCCTGGCGAAGTCGTTAAAGACGTCGATTTCTATGACTATGACGCCAAATATATTGATAATAAAATTACTA
TGGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCGTCAATATGCAAGTAAAGCTTTTAAAGCAATCGGG
GCTTGTGGTTTATCACGCTGTGATTTCTTTTTGACGAAAGATGGACAAATCTTCTTAAACGAACTGAATACAATGCCCGG
TTTTACTCAGTGGTCAATGTATCCTCTGCTTTGGGAAAAT

SEQ ID NO. 1607: SAG0767 FROM THE 18RS21 GBS TYPE II STRAIN (REVERSE COMPLEMENT)

TTGACTTAGCTCTCAAGTATGATAGCCGTATTTTGATTGAACAAGGCGTGACAGCTCGTGAAATTGAAGTAGGTATTTTA
GGCAATAATGATGTTAAGACAACTTTTCCTGGCGAAGTCGTTAAAGACGTCGATTTCTATGACTATGACGCCAAATATAT
TGATAATAAAATTACTATGGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCGTCAATATGCAAGTAAAG
CTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTTCTTTTTGACGAAAGATGGACAAATCTTCTTAAACGAA
CTGAATACAATGCCCGGTTTTACTCAGTGGTCAATGTATCCCCTGCTTTGGGAAAAGTATGGGGCTAACCTT

SEQ ID NO. 1608: SAG0767 FROM THE 18RS21 GBS TYPE II STRAIN

ATCTGTACTGTCTGCAGAAAAGCGTCATGCGTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAA
GTAGGTCAATTTATTAAAACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTAATGACAAACCAAACTGTTGA
TTTAGACAAAATGGTTCGTCCAAGTGATATCTATGATGATAATGCAATTGTTTTCCCCGTTTTACATGGACCAATGGGGG
AAGATGGTTCTATCCAAGGATTTTTAGAAGTTTTAAGGATGCCTTATGTTGGGACTAATATTCTATCTTCAA

Table 16: Comparative Sequences relating to SAG0767 (D-alanine – D-alanine ligase)

SEQ ID NO. 1609: SAG0767 FROM THE 2603V/R GBS TYPE V STRAIN (REVERSE COMPLEMENT)
GGCTATGGATAAAATTACAACAAAACAAGTCCTTGCAACAGTAGGTGTACCTCAGGTTGCATATCAAACTTATTTTGAGG
GTGATGATTTGGAACATGCGATTAAACTCTCTTTAGAAACTTTAAGTTTCCCAATTTTTGTAAAACCGGCTAATATGGGG
TCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCACTTCGCTCTGCAATTGACTTAGCTCTCAAGTATGATAGCCG
TATTTTGATTGAACAAGGCGTGACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAAGACAACTTTTC
CTGGCGAAGTCGTTAAAGACGTCGATTTCTATGACTATGACGCCAAATATATTGATAATAAAATTACTATGGATATTCCA
GCTAAAGTTGATGAAGCAACTATGGAAGCAATGCGTCAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTT
ATCACGCTGTGATTTCTTTTTGACGAAAGAATGGACAAATCTTCTTAAACGAACTGAAATAC

SEQ ID NO. 1610: SAG0767 FROM THE 2603V/R GBS TYPE V STRAIN
TCTGTACTGTCTGCAGAAAGCGTCATGCGTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAAGT
AGGTCAATTTATTAAAACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTAATGACAAACCAAACTGTTGATT
TAGACAAAATGGTTCGTCCAAGTGATATCTATGATGATAAT

SEQ ID NO. 1611: SAG0767 FROM THE H36b GBS TYPE Ib STRAIN (REVERSE COMPLEMENT)
AAAACCGGCTAATATGGGGTCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCACTTCGCTCTGCAATTGACTTAG
CTCTCAAGTATGATAGCCGTATTTTGATTGAACAAGGCGTGACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAAT
GATGTTAAGACAACTTTTCCTGGCGAAGTCGTTAAAGACGTCGATTTCTATGACTATGACGCCAAATATATTGATAATAA
AATTACTATGGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCGTCAATATGCAAGTAAAGCTTTTAAAG
CAATCGGGGCTTGTGGTTTATCACGCTGTGATTTCTTTTTGACGAAAGATGGACAAATCTTCTTAAACGAACTGAATACA
ATGCCCGGTTTTACTCAGTGGTCAATGTATCCCCTGCTTTGGGAAAATATGGGGCTAACTTATAG

SEQ ID NO. 1612: SAG0767 FROM THE H36b TYPE 1b STRAIN
CGTGAAGTATCTGTACTGTCTGCAGAAAGCGTCATGCGTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTAT
CACGCAAGTAGGTCAATTTATTAAAACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTAATGACAAACCAAA
CTGTTGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATGATGATAATGCAATTGTTTTCCCCGTTTTACATGGACCA
ATGGGGGAAGATGGTTCTATCCAAGGATTTTTAGAAGTTTTAAGGATGCCTTATGTTGGGACTAATATTCTATCTTCAAG
CGTGGCTATGGATAAAATTACAACAAAACAAGTCCTTGCAACAGTAG

SEQ ID NO. 1613: SAG0767 FROM THE M732 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
ATGCGATTAAACTCTCTTTAGAACCTTTAAGTTTCCCAATTTTTGTAAACCCGGCTAATATGGGGTCATCAGTAGGTATT
TCAAAAGCGACAGATGAATCCTCACTTCGCTCTGCAATTGACTTAGCTCTCAAGTATGATAGCCGTATTTTGATTGAACA
AGGCGTGACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAAGACAACTTTTCCTGGCGAAGTTGTTA
AAGACGTCGATTTCTATGACTATGACGCCAAATATATTGATAATAAAATTACTATGGATATTCCAGCTAAAGTTGATGAA
GCAACTATGGAAGCAATGCGTCAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTT
CTTTTTGACGAAAGATGGACAAATCTTCTTAAACGAACTGAATACAATGCCCGGTTTTACTCAGTGGTCAATGTATCCTC
TGCTTTGGGAAAATATGGGGCTAACTT

SEQ ID NO. 1614: SAG0767 FROM THE M732 GBS TYPE III STRAIN
GTCATGCCGTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAAGTAGGTCAATTTATTAAAACAC
AAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTAATGACAAACCAAACTGTTGATTTAGACAAAATGGTTCGTCCA
AGTGATATCTATGATGATAATGCAATTGTTTTCCCCGTTTTACATGGACCAATGGGGGAAGATGGTTCTATCCAAGGATT
TTTAGAAGTTTTAAGGATGCCTTATGTTGGGACTAATATTCTATCTTCAAGCGTGGCTATGGATAAAATTACAACAAAAC
AAGTCCTTGCAACAGTAGGTGTACCTCAGG

SEQ ID NO. 1615: SAG0767 FROM THE A909 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
TTTTGAGGGTGATGATTTGGAACATGCGATTAAACTCTCTTTAGAAACTTTAAGTTTCCCAATTTTTGTAAAACCGGCTA
ATATGGGGTCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCACTTCGCTCTGCAATTGACTTAGCTCTCAAGTAT
GATAGCCGTATTTTGATTGAACAAGGCGTGACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAAGAC
AACTTTTCCTGGCGAAGTCGTTAAAGACGTCGATTTCTATGACTATGACGCCAAATATATTGATAATAAAATTACTATGG
ATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCGTCAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCT
TGTGGTTTATCACGCTGTGATTTCTTTTTGACGAAAGATGGACAAATCTTCTTAAACGAACTGAATACAATGCCCGGTTT
TACTCAGTGGTCAATGTATCCCCTGCTTTGGGAAAATATGGGGCTAACTTATAGTGA

SEQ ID NO. 1616: SAG0767 FROM THE A909 GBS TYPE Ia STRAIN
TGGTCGCTCTGCGGAACGTGAAGTATCTGTACTGTCTGCAGAAAGCGTCATGCGTGCTATTAATTATGATAAATTTTTTG
TTAAAACTTATTTTATCACGCAAGTAGGTCAATTTATTAAAACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAG
TTAATGACAAACCAAACTGTTGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATGATGATAATGCAATTGTTTTCCC
CGTTTTACATGGACCAATGGGGGAAGATGGTTCTATCCAAGGATTTTTAGAAGTTTTAAGGATGCCTTATGTTGGGACTA
ATATTCTATCTTCAAGCGTGGCTATGGATAAAATTACAACAAAACAAGTCCTTGCAACAGTAGG

Table 16: Comparative Sequences relating to SAG0767 (D-alanine – D-alanine ligase)

SEQ ID NO. 1617: SAG0767 FROM THE JM9130013 GBS TYPE VIII STRAIN (REVERSE COMPLEMENT)

AAGCAGGGGATACATTGACCACTGAGTAAAACCGGGCATTGTATTCAGTTCGTTTAAGAAGATCTGTCCATCTTTCGTCA
AAAAGAAATCACAGCGTGATAAACCACAAGCCCCGATTGCTTTAAAAGCTTTACTTGCATATTGACGCATTGCTTCCATA
GATGCTTCATCAACTTTAGCTGGAATATCCATAGCAATTTTATTATCAATATATTTGGCG

```
SEQ1601         GGTCGCTCTGTCGGAACGTGAAGTATCTGTACTGTCTGCAGAAAGCGTCATGCGTGCTA
SEQ1602         -----------------------------------------------------------
SEQ1603         -----------------------------------------------------------
SEQ1604         -----------------------------------------------------------
SEQ1605         -----------------------------------------------------------
SEQ1606         -----------------------------------------------------------
SEQ1607         -----------------------------------------------------------
SEQ1608         -----------------------------------------------------------
SEQ1609         -----------------------------------------------------------
SEQ1610         -----------------------------------------------------------
SEQ1611         -----------------------------------------------------------
SEQ1612         -----------------------------------------------------------
SEQ1613         -----------------------------------------------------------
SEQ1614         -----------------------------------------------------------
SEQ1615         -----------------------------------------------------------
SEQ1616         -----------------------------------------------------------
SEQ1617         -----------------------------------------------------------

SEQ1601         TAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAAGTAGGTCAATTTATTA
SEQ1602         -----------------------------------------------------------
SEQ1603         -----------------------------------------------------------
SEQ1604         -----------------------------------------------------------
SEQ1605         -----------------------------------------------------------
SEQ1606         -----------------------------------------------------------
SEQ1607         -----------------------------------------------------------
SEQ1608         -----------------------------------------------------------
SEQ1609         -----------------------------------------------------------
SEQ1610         -----------------------------------------------------------
SEQ1611         -----------------------------------------------------------
SEQ1612         -----------------------------------------------------------
SEQ1613         -----------------------------------------------------------
SEQ1614         -----------------------------------------------------------
SEQ1615         -----------------------------------------------------------
SEQ1616         -----------------------------------------------------------
SEQ1617         -----------------------------------------------------------

SEQ1601         AACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTAATGACAAACCAAACTG
SEQ1602         -----------------------------------------------------------
SEQ1603         -----------------------------------------------------------
SEQ1604         -----------------------------------------------------------
SEQ1605         -----------------------------------------------------------
SEQ1606         -----------------------------------------------------------
SEQ1607         -----------------------------------------------------------
SEQ1608         -----------------------------------------------------------
SEQ1609         -----------------------------------------------------------
SEQ1610         -----------------------------------------------------------
SEQ1611         -----------------------------------------------------------
SEQ1612         -----------------------------------------------------------
SEQ1613         -----------------------------------------------------------
SEQ1614         -----------------------------------------------------------
SEQ1615         -----------------------------------------------------------
SEQ1616         -----------------------------------------------------------
SEQ1617         -----------------------------------------------------------
```

Table 16: Comparative Sequences relating to SAG0767 (D-alanine – D-alanine ligase)

```
SEQ1601         TGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATGATGATAATGCAATTGTTTTCC
SEQ1602         -----------------------------------------------------------
SEQ1603         -----------------------------------------------------------
SEQ1604         -----------------------------------------------------------
SEQ1605         -----------------------------------------------------------
SEQ1606         -----------------------------------------------------------
SEQ1607         -----------------------------------------------------------
SEQ1608         -----------------------------------------------------------
SEQ1609         -----------------------------------------------------------
SEQ1610         -----------------------------------------------------------
SEQ1611         -----------------------------------------------------------
SEQ1612         -----------------------------------------------------------
SEQ1613         -----------------------------------------------------------
SEQ1614         -----------------------------------------------------------
SEQ1615         -----------------------------------------------------------
SEQ1616         -----------------------------------------------------------
SEQ1617         -----------------------------------------------------------

SEQ1601         CGTTTTACATGGACCAATGGGGGAAGATGGTTCTATCCAAGGATTTTTAGAAGTTTTAA
SEQ1602         -----------------------------------------------------------
SEQ1603         -----------------------------------------------------------
SEQ1604         -----------------------------------------------------------
SEQ1605         -----------------------------------------------------------
SEQ1606         -----------------------------------------------------------
SEQ1607         -----------------------------------------------------------
SEQ1608         -----------------------------------------------------------
SEQ1609         -----------------------------------------------------------
SEQ1610         -----------------------------------------------------------
SEQ1611         -----------------------------------------------------------
SEQ1612         -----------------------------------------------------------
SEQ1613         -----------------------------------------------------------
SEQ1614         -----------------------------------------------------------
SEQ1615         -----------------------------------------------------------
SEQ1616         -----------------------------------------------------------
SEQ1617         -----------------------------------------------------------

SEQ1601         GATGCCTTATGTTGGGACTAATATTCTATCTTCAAGCGTGGCTATGGATAAAATTACAA
SEQ1602         -----------------------------------------------------------
SEQ1603         -----------------------------------------------------------
SEQ1604         -----------------------------------------------------------
SEQ1605         -----------------------------------------------------------
SEQ1606         -----------------------------------------------------------
SEQ1607         -----------------------------------------------------------
SEQ1608         -----------------------------------------------------------
SEQ1609         ---------------------------------------GGCTATGGATAAAATTACAA
SEQ1610         -----------------------------------------------------------
SEQ1611         -----------------------------------------------------------
SEQ1612         -----------------------------------------------------------
SEQ1613         -----------------------------------------------------------
SEQ1614         -----------------------------------------------------------
SEQ1615         -----------------------------------------------------------
SEQ1616         -----------------------------------------------------------
SEQ1617         -----------------------------------------------------------
```

Table 16: Comparative Sequences relating to SAG0767 (D-alanine – D-alanine ligase)

```
SEQ1601     AAAACAAGTCCTTGCAACAGTAGGTGTACCTCAGGTTGCATATCAAACTTATTTTGAGG
SEQ1602     -----------------------------------------------------------
SEQ1603     -----------------------------------------------------------
SEQ1604     -----------------------------------------------------------
SEQ1605     -----------------------------------------------------------
SEQ1606     -----------------------------------------------------------
SEQ1607     -----------------------------------------------------------
SEQ1608     -----------------------------------------------------------
SEQ1609     AAAACAAGTCCTTGCAACAGTAGGTGTACCTCAGGTTGCATATCAAACTTATTTTGAGG
SEQ1610     -----------------------------------------------------------
SEQ1611     -----------------------------------------------------------
SEQ1612     -----------------------------------------------------------
SEQ1613     -----------------------------------------------------------
SEQ1614     -----------------------------------------------------------
SEQ1615     ---------------------------------------------------TTTTGAGG
SEQ1616     -----------------------------------------------------------
SEQ1617     -----------------------------------------------------------

SEQ1601     TGATGATTTGGAACATGCGATTAAACTCTCTTTAGAAACTTTAAGTTTCCCAATTTTTG
SEQ1602     -----------------------------------------------------------
SEQ1603     -----------------------------------------------------------
SEQ1604     -----------------------------------------------------------
SEQ1605     -----------------------------------------------------------
SEQ1606     -----------------------------------------------------------
SEQ1607     -----------------------------------------------------------
SEQ1608     -----------------------------------------------------------
SEQ1609     TGATGATTTGGAACATGCGATTAAACTCTCTTTAGAAACTTTAAGTTTCCCAATTTTTG
SEQ1610     -----------------------------------------------------------
SEQ1611     -----------------------------------------------------------
SEQ1612     -----------------------------------------------------------
SEQ1613     --------------ATGCGATTAAACTCTCTTTAGAACCTTTAAGTTTCCCAATTTTTG
SEQ1614     -----------------------------------------------------------
SEQ1615     TGATGATTTGGAACATGCGATTAAACTCTCTTTAGAAACTTTAAGTTTCCCAATTTTTG
SEQ1616     -----------------------------------------------------------
SEQ1617     -----------------------------------------------------------

SEQ1601     AAAACCGGCTAATATGGGGTCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCAC
SEQ1602     ---------------------------------------------------AAACCGGC
SEQ1603     ------------TCGCTCTGCGGAACGTGAAGTATCTGTACTG-TCTGCAGAAA-GCGT
SEQ1604     -----------------------------------------------------------
SEQ1605     -----------------------AACGTGAAGTATCTGTACTGCTCTGCAGAAAAGCGT
SEQ1606     --------CTAATATGGGGTCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCAC
SEQ1607     -----------------------------------------------------------
SEQ1608     ---------------------------------ATCTGTACTG-TCTGCAGAAAAGCGT
SEQ1609     AAAACCGGCTAATATGGGGTCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCAC
SEQ1610     ----------------------------------TCTGTACTG-TCTGCAGAAA-GCGT
SEQ1611     AAAACCGGCTAATATGGGGTCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCAC
SEQ1612     -------------------------CGTGAAGTATCTGTACTG-TCTGCAGAAA-GCGT
SEQ1613     AAACCCGGCTAATATGGGGTCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCAC
SEQ1614     ---------------------------------------------------------GT
SEQ1615     AAAACCGGCTAATATGGGGTCATCAGTAGGTATTTCAAAAGCGACAGATGAATCCTCAC
SEQ1616     ---------TGGTCGCTCTGCGGAACGTGAAGTATCTGTACTG-TCTGCAGAAA-GCGT
SEQ1617     ----------------------AAGCAGGGGATACATTGACCACTGAGTAAAACCGGGC

SEQ1601     TCGCTCTGCAATTGACTTAGCTCTCAAGTATGATAGCCGTATTTTGATTGAACAAGGCG
SEQ1602     TTGT-ATTCAGTTCGTTTAAGAAGACTTGTCCATCTTTCGTCAAAAAGAAATCACAGCG
SEQ1603     ATGC-GTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAAGTAG
SEQ1604     -----------------------------------------------------------
SEQ1605     ATGC-GTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAAGTAG
SEQ1606     TCGCTCTGCAATTGACTTAGCTCTCAAGTATGATAGCCGTATTTTGATTGAACAAGGCG
SEQ1607     -----------TTGACTTAGCTCTCAAGTATGATAGCCGTATTTTGATTGAACAAGGCG
SEQ1608     ATGC-GTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAAGTAG
SEQ1609     TCGCTCTGCAATTGACTTAGCTCTCAAGTATGATAGCCGTATTTTGATTGAACAAGGCG
```

Table 16: Comparative Sequences relating to SAG0767 (D-alanine – D-alanine ligase)

```
SEQ1610          ATGC-GTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAAGTAG
SEQ1611          TCGCTCTGCAATTGACTTAGCTCTCAAGTATGATAGCCGTATTTTGATTGAACAAGGCG
SEQ1612          ATGC-GTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAAGTAG
SEQ1613          TCGCTCTGCAATTGACTTAGCTCTCAAGTATGATAGCCGTATTTTGATTGAACAAGGCG
SEQ1614          ATGCCGTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAAGTAG
SEQ1615          TCGCTCTGCAATTGACTTAGCTCTCAAGTATGATAGCCGTATTTTGATTGAACAAGGCG
SEQ1616          ATGC-GTGCTATTAATTATGATAAATTTTTTGTTAAAACTTATTTTATCACGCAAGTAG
SEQ1617          TTGT-ATTCAGTTCGTTTAAGAAGATCTGTCCATCTTTCGTCAAAAAGAAATCACAGCG

SEQ1601          GACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAAGACAACTTTTC
SEQ1602          GATAAACCACAAGC----CCCGATTGCTTTAAAAGCTTTACTTGCATATTGACGCATTG
SEQ1603          GTCAATTTATTAAA---ACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTA
SEQ1604          -----------------------------------------------------------
SEQ1605          GTCAATTTATTAAA---ACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAA----
SEQ1606          GACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAAGACAACTTTTC
SEQ1607          GACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAAGACAACTTTTC
SEQ1608          GTCAATTTATTAAA---ACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTA
SEQ1609          GACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAAGACAACTTTTC
SEQ1610          GTCAATTTATTAAA---ACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTA
SEQ1611          GACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAAGACAACTTTTC
SEQ1612          GTCAATTTATTAAA---ACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTA
SEQ1613          GACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAAGACAACTTTTC
SEQ1614          GTCAATTTATTAAA---ACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTA
SEQ1615          GACAGCTCGTGAAATTGAAGTAGGTATTTTAGGCAATAATGATGTTAAGACAACTTTTC
SEQ1616          GTCAATTTATTAAA---ACACAAGAATTTGATGAAATGCCATCTTCAGATGAAAAGTTA
SEQ1617          GATAAACCACAAGC----CCCGATTGCTTTAAAAGCTTTACTTGCATATTGACGCATTG

SEQ1601          TGGCGAAGTTGTTAAAGACGTCGATTTCTATGA--CTATGACGCCAAAT-ATATTGATA
SEQ1602          TTCCATAGTT----GCTTCATCAACTTTAGCTGGAATATCCATAGTAATTTTATTATCA
SEQ1603          TGACAAACC-----AAACTGTTGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATG
SEQ1604          ------------------CGTCGATTTCTATGA--CTATGACGCCAAAT-ATATTGATA
SEQ1605          -----------------------------------------------------------
SEQ1606          TGGCGAAGTCGTTAAAGACGTCGATTTCTATGA--CTATGACGCCAAAT-ATATTGATA
SEQ1607          TGGCGAAGTCGTTAAAGACGTCGATTTCTATGA--CTATGACGCCAAAT-ATATTGATA
SEQ1608          TGACAAACC-----AAACTGTTGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATG
SEQ1609          TGGCGAAGTCGTTAAAGACGTCGATTTCTATGA--CTATGACGCCAAAT-ATATTGATA
SEQ1610          TGACAAACC-----AAACTGTTGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATG
SEQ1611          TGGCGAAGTCGTTAAAGACGTCGATTTCTATGA--CTATGACGCCAAAT-ATATTGATA
SEQ1612          TGACAAACC-----AAACTGTTGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATG
SEQ1613          TGGCGAAGTTGTTAAAGACGTCGATTTCTATGA--CTATGACGCCAAAT-ATATTGATA
SEQ1614          TGACAAACC-----AAACTGTTGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATG
SEQ1615          TGGCGAAGTCGTTAAAGACGTCGATTTCTATGA--CTATGACGCCAAAT-ATATTGATA
SEQ1616          TGACAAACC-----AAACTGTTGATTTAGACAAAATGGTTCGTCCAAGTGATATCTATG
SEQ1617          TTCCATAGAT----GCTTCATCAACTTTAGCTGGAATATCCATAGCAATTTTATTATCA

SEQ1601          TAAAATTACTAT--GGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCG
SEQ1602          TATATTTGGCGTCATAGTCATAGAAATCGACGTCTTTAACGACTTCGCCAGG--AAAAG
SEQ1603          TGATAATGCAAT--TGTTTTCCCCGTTTTAC---ATGGACCAATGGGGGAAG--ATGGT
SEQ1604          TAAAATTACTAT--GGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCG
SEQ1605          -----------------------------------------------------------
SEQ1606          TAAAATTACTAT--GGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCG
SEQ1607          TAAAATTACTAT--GGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCG
SEQ1608          TGATAATGCAAT--TGTTTTCCCCGTTTTAC---ATGGACCAATGGGGGAAG--ATGGT
SEQ1609          TAAAATTACTAT--GGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCG
SEQ1610          TGATAAT----------------------------------------------------
SEQ1611          TAAAATTACTAT--GGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCG
SEQ1612          TGATAATGCAAT--TGTTTTCCCCGTTTTAC---ATGGACCAATGGGGGAAG--ATGGT
SEQ1613          TAAAATTACTAT--GGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCG
SEQ1614          TGATAATGCAAT--TGTTTTCCCCGTTTTAC---ATGGACCAATGGGGGAAG--ATGGT
SEQ1615          TAAAATTACTAT--GGATATTCCAGCTAAAGTTGATGAAGCAACTATGGAAGCAATGCG
SEQ1616          TGATAATGCAAT--TGTTTTCCCCGTTTTAC---ATGGACCAATGGGGGAAG--ATGGT
SEQ1617          TATATTTGGCGT----ABLECMPARATIVESEQENCESRELA-TINGTSAGD--ALANI
```

Table 16: Comparative Sequences relating to SAG0767 (D-alanine – D-alanine ligase)

```
SEQ1601         CAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTT
SEQ1602         T--GTCTTAACATCATTATTGCCTAAAATACCTACTTCAATTTCACGAGCTGTCACGCC
SEQ1603         C--TATCCAAG--GATTTTTAGAAGTTTTAAGGATGCCTTATGTTGGGACTAATATTCT
SEQ1604         CAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTT
SEQ1605         -----------------------------------------------------------
SEQ1606         CAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTT
SEQ1607         CAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTT
SEQ1608         C--TATCCAAG--GATTTTTAGAAGTTTTAAGGATGCCTTATGTTGGGACTAATATTCT
SEQ1609         CAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTT
SEQ1610         -----------------------------------------------------------
SEQ1611         CAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTT
SEQ1612         C--TATCCAAG--GATTTTTAGAAGTTTTAAGGATGCCTTATGTTGGGACTAATATTCT
SEQ1613         CAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTT
SEQ1614         C--TATCCAAG--GATTTTTAGAAGTTTTAAGGATGCCTTATGTTGGGACTAATATTCT
SEQ1615         CAATATGCAAGTAAAGCTTTTAAAGCAATCGGGGCTTGTGGTTTATCACGCTGTGATTT
SEQ1616         C--TATCCAAG--GATTTTTAGAAGTTTTAAGGATGCCTTATGTTGGGACTAATATTCT
SEQ1617         E--DALANINELIGASE------------------------------------------

SEQ1601         TTTTTGACGAAAGA--TGGACAAATCTTCTTAAACGAACTGAA-TACAATGCCCGGTTT
SEQ1602         TGTTCAATCAAAATACGGCTATCATACTTGAGAGCTAAGTCAATKSCAGAGCGAAGTGA
SEQ1603         TCTTCAAGCGTGGCTAT------------------------------------------
SEQ1604         TTTTTGACGAAAGA--TGGACAAATCTTCTTAAACGAACTGAA-TACAATGCCC-----
SEQ1605         -----------------------------------------------------------
SEQ1606         TTTTTGACGAAAGA--TGGACAAATCTTCTTAAACGAACTGAA-TACAATGCCCGGTTT
SEQ1607         TTTTTGACGAAAGA--TGGACAAATCTTCTTAAACGAACTGAA-TACAATGCCCGGTTT
SEQ1608         TCTTCAA----------------------------------------------------
SEQ1609         TTTTTGACGAAAGAA-TGGACAAATCTTCTTAAACGAACTGAAATAC------------
SEQ1610         -----------------------------------------------------------
SEQ1611         TTTTTGACGAAAGA--TGGACAAATCTTCTTAAACGAACTGAA-TACAATGCCCGGTTT
SEQ1612         TCTTCAAGCGTGGCTATGGATAAAATTACAACAAAACAAGTCCTTGCAACAGTAG----
SEQ1613         TTTTTGACGAAAGA--TGGACAAATCTTCTTAAACGAACTGAA-TACAATGCCCGGTTT
SEQ1614         TCTTCAAGCGTGGCTATGGATAAAATTACAACAAAACAAGTCCTTGCAACAGTAGGTGT
SEQ1615         TTTTTGACGAAAGA--TGGACAAATCTTCTTAAACGAACTGAA-TACAATGCCCGGTTT
SEQ1616         TCTTCAAGCGTGGCTATGGATAAAATTACAACAAAACAAGTCCTTGCAACAGTAGG---
SEQ1617         -----------------------------------------------------------

SEQ1601         ACTCAGTGGTCAATGTATCCTCTGCTTTGGGAAAA-TATGGGGCTAACTTATAGTGATT
SEQ1602         GATTCATCTGTCGCTTTTGAAATACCTACTGATGACCCCATATTAGCCGGTTTTACAAA
SEQ1603         -----------------------------------------------------------
SEQ1604         -----------------------------------------------------------
SEQ1605         -----------------------------------------------------------
SEQ1606         ACTCAGTGGTCAATGTATCCTCTGCTTTGGGAAAA-T----------------------
SEQ1607         ACTCAGTGGTCAATGTATCCCCTGCTTTGGGAAAAGTATGGGGCTAACCTT--------
SEQ1608         -----------------------------------------------------------
SEQ1609         -----------------------------------------------------------
SEQ1610         -----------------------------------------------------------
SEQ1611         ACTCAGTGGTCAATGTATCCCCTGCTTTGGGAAAA-TATGGGGCTAACTTATAG-----
SEQ1612         -----------------------------------------------------------
SEQ1613         ACTCAGTGGTCAATGTATCCTCTGCTTTGGGAAAA-TATGGGGCTAACTT---------
SEQ1614         CCTCAGG----------------------------------------------------
SEQ1615         ACTCAGTGGTCAATGTATCCCCTGCTTTGGGAAAA-TATGGGGCTAACTTATAGTGA--
SEQ1616         -----------------------------------------------------------
SEQ1617         -----------------------------------------------------------
```

Table 16: Comparative Sequences relating to SAG0767 (D-alanine – D-alanine ligase)

```
SEQ1601         GATTG------------------------------------------------------
SEQ1602         ATTGGGAAACTTAAAGTTTCTAAAGAGAGTTTAATCGCATGTTCCAAATCATCACCCTC
SEQ1603         -----------------------------------------------------------
SEQ1604         -----------------------------------------------------------
SEQ1605         -----------------------------------------------------------
SEQ1606         -----------------------------------------------------------
SEQ1607         -----------------------------------------------------------
SEQ1608         -----------------------------------------------------------
SEQ1609         -----------------------------------------------------------
SEQ1610         -----------------------------------------------------------
SEQ1611         -----------------------------------------------------------
SEQ1612         -----------------------------------------------------------
SEQ1613         -----------------------------------------------------------
SEQ1614         -----------------------------------------------------------
SEQ1615         -----------------------------------------------------------
SEQ1616         -----------------------------------------------------------
SEQ1617         -----------------------------------------------------------

SEQ1601         -----------------------------------------------------------
SEQ1602         AAATAAGTTTGATATGCAACCTGAGGTACACCTACTGTTGCAAGGACTTGTTTTGTTGT
SEQ1603         -----------------------------------------------------------
SEQ1604         -----------------------------------------------------------
SEQ1605         -----------------------------------------------------------
SEQ1606         -----------------------------------------------------------
SEQ1607         -----------------------------------------------------------
SEQ1608         -----------------------------------------------------------
SEQ1609         -----------------------------------------------------------
SEQ1610         -----------------------------------------------------------
SEQ1611         -----------------------------------------------------------
SEQ1612         -----------------------------------------------------------
SEQ1613         -----------------------------------------------------------
SEQ1614         -----------------------------------------------------------
SEQ1615         -----------------------------------------------------------
SEQ1616         -----------------------------------------------------------
SEQ1617         -----------------------------------------------------------

SEQ1601         -----------------------------------------------------------
SEQ1602         ATTTTATCCATAGCCACGCTTGAAGATAGAATATTAGTCCCAACATAAGGCATCCTTAA
SEQ1603         -----------------------------------------------------------
SEQ1604         -----------------------------------------------------------
SEQ1605         -----------------------------------------------------------
SEQ1606         -----------------------------------------------------------
SEQ1607         -----------------------------------------------------------
SEQ1608         -----------------------------------------------------------
SEQ1609         -----------------------------------------------------------
SEQ1610         -----------------------------------------------------------
SEQ1611         -----------------------------------------------------------
SEQ1612         -----------------------------------------------------------
SEQ1613         -----------------------------------------------------------
SEQ1614         -----------------------------------------------------------
SEQ1615         -----------------------------------------------------------
SEQ1616         -----------------------------------------------------------
SEQ1617         -----------------------------------------------------------
```

Table 16: Comparative Sequences relating to SAG0767 (D-alanine – D-alanine ligase)

```
SEQ1601          ------------------------------------------------------------
SEQ1602          ACTTCTAAAAATCCTTGGATAGAACCATCTTCCCCCATTGGTCCATGTAAAACGGGGAA
SEQ1603          ------------------------------------------------------------
SEQ1604          ------------------------------------------------------------
SEQ1605          ------------------------------------------------------------
SEQ1606          ------------------------------------------------------------
SEQ1607          ------------------------------------------------------------
SEQ1608          ------------------------------------------------------------
SEQ1609          ------------------------------------------------------------
SEQ1610          ------------------------------------------------------------
SEQ1611          ------------------------------------------------------------
SEQ1612          ------------------------------------------------------------
SEQ1613          ------------------------------------------------------------
SEQ1614          ------------------------------------------------------------
SEQ1615          ------------------------------------------------------------
SEQ1616          ------------------------------------------------------------
SEQ1617          ------------------------------------------------------------

SEQ1601          ------------------------------------------------------------
SEQ1602          ACAATTGCATTATCATCATAGATATCACTTGGACGAACCATTTTGTCTAAATCAACAGT
SEQ1603          ------------------------------------------------------------
SEQ1604          ------------------------------------------------------------
SEQ1605          ------------------------------------------------------------
SEQ1606          ------------------------------------------------------------
SEQ1607          ------------------------------------------------------------
SEQ1608          ------------------------------------------------------------
SEQ1609          ------------------------------------------------------------
SEQ1610          ------------------------------------------------------------
SEQ1611          ------------------------------------------------------------
SEQ1612          ------------------------------------------------------------
SEQ1613          ------------------------------------------------------------
SEQ1614          ------------------------------------------------------------
SEQ1615          ------------------------------------------------------------
SEQ1616          ------------------------------------------------------------
SEQ1617          ------------------------------------------------------------

SEQ1601          ------------------------------------------------------------
SEQ1602          TGGTTTGTCATTAACTTTTCATCTGAAGATGGCATTTCATCAAATTCTTGTGTTTTAAT
SEQ1603          ------------------------------------------------------------
SEQ1604          ------------------------------------------------------------
SEQ1605          ------------------------------------------------------------
SEQ1606          ------------------------------------------------------------
SEQ1607          ------------------------------------------------------------
SEQ1608          ------------------------------------------------------------
SEQ1609          ------------------------------------------------------------
SEQ1610          ------------------------------------------------------------
SEQ1611          ------------------------------------------------------------
SEQ1612          ------------------------------------------------------------
SEQ1613          ------------------------------------------------------------
SEQ1614          ------------------------------------------------------------
SEQ1615          ------------------------------------------------------------
SEQ1616          ------------------------------------------------------------
SEQ1617          ------------------------------------------------------------
```

Table 16: Comparative Sequences relating to SAG0767 (D-alanine – D-alanine ligase)

```
SEQ1601          ------------------
SEQ1602          AATTGACCTACTTGCGTG
SEQ1603          ------------------
SEQ1604          ------------------
SEQ1605          ------------------
SEQ1606          ------------------
SEQ1607          ------------------
SEQ1608          ------------------
SEQ1609          ------------------
SEQ1610          ------------------
SEQ1611          ------------------
SEQ1612          ------------------
SEQ1613          ------------------
SEQ1614          ------------------
SEQ1615          ------------------
SEQ1616          ------------------
SEQ1617          ------------------
```

Table 17: Comparative Sequences relating to SAG1086 (xanthine phophoribosyltransferase)

SEQ ID NO. 1701: SAG1086 FROM THE1169NT1 GBS NONTYPEABLE STRAIN
TTTAAAGGTTGATTCCTTTTTGACTCATCAGGTAGATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATAAATATAAAGAAG
CCGGCATTACGAAGGTTGTTACGATTGAAGCATCTGGAATTGCGCCAGCAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTT
GCTAAAAAGGCTAAGAACATTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGWTACGAGTCAAGTTTC
TATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATGACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTG
AAATTATTGGTCAAGCTGGAGCTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTGTTAGAAAAA
ACAGGTGTTCCAGT

SEQ ID NO. 1702: SAG0767 FROM THE 18RS21 GBS TYPE II STRAIN
TTTAGGTGAGAACATTTTAAAGGTTGATTCTTTTTTGACTCATCAGGTAGATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTG
ATAAATATAAAGAAGCCGGCATTACGAAGGTTGTTACGATTGAAGCATCTGGAATTGCACCAGCAGTGTACGCAGCTCAAGCATTGGGC
GkACCAATGATATTTGCTAAAAAAGCTAAGAACATTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGT
TACGAGTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATGACTTTTTAGCAAACGGTCAAGCGG
CTAAAGGATTACTTGAAATTATTGGTCAAGCTGGAGCTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGT
GATTTGTTAGAAAAAACA

SEQ ID NO. 1703: SAG0767 FROM THE H36b1 GBS TYPE Ib STRAIN
AAGAACGTATTCTTAAAGATGGTGATGTTTTAGGTGAGAACATTTTAAAAGTTGATTCTTTTTTGACTCATCAGGTAGATTTTGAGTTA
ATGCAGGAAATAGGTAAAGTTTTTGCTGATAAATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCC
AGCAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACATTACTATGACTGAAGGTATCTTAACTG
CTGAAGTGTATTCTTTTACAAAGCAAGTTACGAGTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATT
GATGACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATTATTGGTCAAGCTGGAGCTAAGGTTGCTGGTATCGGAATCYT
TATTGAAAAATCTTTCCAAGATGGGCGTGATT

SEQ ID NO. 1704: SAG0767 FROM THE M732 GBS TYPE III STRAIN
ATTCTTTTTTGACTATCAGGTAAATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATAAATATAAAGAAGCCGGCATTACGA
AGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAGCAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCT
AAGAACATTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGAGTCAAGTTTCTATTGTGAGTCG
CTTTTTATCTAACGATGATACTGTACTCATCATTGATGACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATTATTGGTC
AAGCTGAAGCTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTGTTAGAAAAAACAGGTGTTCCG
GTTACTTCTCTTGCTCGT

SEQ ID NO. 1705: SAG0767 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
GAACGTATTCTTAAAGATGGTGATGTTTTAGGTGAGAACATTTTAAAAGTTGATTCTTTTTTGACTCATCAGGTAAATTTTGAGTTAAT
GCAGGAAATAGGTAAAGTTTTTGCTGATAAATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAG
CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACATTACTATGACTGAAGGTATRTTAACTGCT
GAAGTGTATTCTTTTACAAAGCAAGTTACGAGTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGA
TGACTTTTTAACAAACGGTCAAGC

SEQ ID NO. 1706: SAG0767 FROM THE 090 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
ACATTTTAAAGGTTGATTCTTTTTTGACTCATCAGGTAGATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATAAATATAAA
GAAGCCGGCATTACGAAGGTTGTTACGATTGAAGCATCTGGAATTGCACCAGCAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGAT
ATTTGCTAAAAAAGCTAAGAACATTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGAGTCAAG
TTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATGACTTTTTAGCAAACMGTCYAGCGGCTAAAGGATTA
CTTGAAATTATTGGTCAAGCTGGAGCTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTGTTAGA
AAA

SEQ ID NO. 1707: SAG0767 FROM THE A909 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
ACGTATTCTTAAAGATGGTGATGTTTTAGGTGAGAACATTTTAAAAGTTGATTCTTTTTTGACTCATCAGGTAGATTTTGAGTTAATGC
AGGAAATAGGTAAAGTTTTTGCTGATAAATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAGCA
GTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACATTACTATGACTGAAGGTATCTTAACTGCTGA
AGTGTATTCTTTTACAAAGCAAGTTACGAGTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
ACTTTTTAGCAAACGGKCAAGCGGSTAAAGGATTACTTGAAATTATTGGTCAAGCTGGAGCTA

SEQ ID NO. 1708: SAG0767 FROM THE COH1 GBS TYPE Ia STRAIN
TTTAAAAGTTGATTCTTTTTTGACTCATCAGGTAAATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATAAATATAAAGAAG
CCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAGCAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTT
GCTAAAAAAGCTAAGAACATTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGAGTCAAGTTTC
TATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATGACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTG
AAATTATTGGTCAAGCTGAAGCTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTGTTAGAAAAA
ACAGGTGTTCCGGTTAC

SEQ ID NO. 1709: SAG0767 FROM THE CJB110 GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)
GCTGATAAATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAGCAGTGTACGCAGCTCAAGCATT
GGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACATTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGC
AAGTTACGAGTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATGACTTTTTAGCAAACGGTCAA
GCGGCTAAAGGATTACTTGAAATTTATTGGTCAAGCTGGAGCTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATG
GGCGTGATTTGTTAGAAAAAACAGGTGTTCCAGT

Table 17: Comparative Sequences relating to SAG1086 (xanthine phophoribosyltransferase)

SEQ ID NO. 1710: SAG0767 FROM THE 2603 V/R GBS TYPE V STRAIN

```
AACGTATTCTTAAAGATGGTGATGTTTTAGGTGAGAACATTTTAAAAGTTGATTCTTTTTGACTCATCAGGTAGATTTTGAGTTAATG
CAGGAAATAGGTAAAGTTTTTGCTGATAAATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAGC
AGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACATTACTATGACTGAAGGTATCTTAACTGCTG
AAGTGTATTCTTTTACAAAGCAAGTTACGAGTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGAT
GACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATTATTGGTCAAGCTGGAGCTAAGGTTGCTGGTATCGGAATCGTTAT
TGAAAAATCTTTCCAAGATGGGCGTGATTTGTTAGAAAAAACAGGTGTTCCAG
```

SEQ ID NO. 1711: SAG0767 FROM THE JM9130013 GBS TYPE VIII STRAIN (REVERSE COMPLEMENT)

```
ACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAGCAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAA
AGCTAAGAACATTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGAGTCAAGTTTCTATTGTGA
GTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATGACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATTATT
GGTCAAGCTGGAGCTAAGGTTGCTGGTATCGGA
```

```
SEQ1701         ------------------------------------------TTTAAAGGTTGATTCCT
SEQ1702         ---------------------------TTTAGGTGAGAACATTTTAAAGGTTGATTCTT
SEQ1703         AGAACGTATTCTTAAAGATGGTGATGTTTTAGGTGAGAACATTTTAAAAGTTGATTCTT
SEQ1704         ------------------------------------------------------ATTCT
SEQ1705         -GAACGTATTCTTAAAGATGGTGATGTTTTAGGTGAGAACATTTTAAAAGTTGATTCTT
SEQ1706         --------------------------------------ACATTTTAAAGGTTGATTCTT
SEQ1707         ---ACGTATTCTTAAAGATGGTGATGTTTTAGGTGAGAACATTTTAAAAGTTGATTCTT
SEQ1708         ------------------------------------------TTTAAAAGTTGATTCTT
SEQ1709         -----------------------------------------------------------
SEQ1710         --AACGTATTCTTAAAGATGGTGATGTTTTAGGTGAGAACATTTTAAAAGTTGATTCTT
SEQ1711         -----------------------------------------------------------

SEQ1701         TTTGACTCATCAGGTAGATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATA
SEQ1702         TTTGACTCATCAGGTAGATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATA
SEQ1703         TTTGACTCATCAGGTAGATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATA
SEQ1704         TTTTGACTATCAGGTAAATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATA
SEQ1705         TTTGACTCATCAGGTAAATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATA
SEQ1706         TTTGACTCATCAGGTAGATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATA
SEQ1707         TTTGACTCATCAGGTAGATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATA
SEQ1708         TTTGACTCATCAGGTAAATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATA
SEQ1709         ----------------------------------------------------GCTGATA
SEQ1710         TTTGACTCATCAGGTAGATTTTGAGTTAATGCAGGAAATAGGTAAAGTTTTTGCTGATA
SEQ1711         -----------------------------------------------------------

SEQ1701         ATATAAAGAAGCCGGCATTACGAAGGTTGTTACGATTGAAGCATCTGGAATTGCGCCAG
SEQ1702         ATATAAAGAAGCCGGCATTACGAAGGTTGTTACGATTGAAGCATCTGGAATTGCACCAG
SEQ1703         ATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAG
SEQ1704         ATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAG
SEQ1705         ATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAG
SEQ1706         ATATAAAGAAGCCGGCATTACGAAGGTTGTTACGATTGAAGCATCTGGAATTGCACCAG
SEQ1707         ATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAG
SEQ1708         ATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAG
SEQ1709         ATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAG
SEQ1710         ATATAAAGAAGCCGGCATTACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAG
SEQ1711         -------------------ACGAAGGTTGTTACAATTGAAGCATCTGGAATTGCGCCAG

SEQ1701         CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAGGCTAAGAACA
SEQ1702         CAGTGTACGCAGCTCAAGCATTGGGCGKACCAATGATATTTGCTAAAAAAGCTAAGAACA
SEQ1703         CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACA
SEQ1704         CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACA
SEQ1705         CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACA
SEQ1706         CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACA
SEQ1707         CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACA
SEQ1708         CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACA
SEQ1709         CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACA
SEQ1710         CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACA
SEQ1711         CAGTGTACGCAGCTCAAGCATTGGGCGTACCAATGATATTTGCTAAAAAAGCTAAGAACA
```

Table 17: Comparative Sequences relating to SAG1086 (xanthine phophoribosyltransferase)

```
SEQ1701        TTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGWTACGA
SEQ1702        TTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGA
SEQ1703        TTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGA
SEQ1704        TTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGA
SEQ1705        TTACTATGACTGAAGGTATRTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGA
SEQ1706        TTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGA
SEQ1707        TTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGA
SEQ1708        TTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGA
SEQ1709        TTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGA
SEQ1710        TTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGA
SEQ1711        TTACTATGACTGAAGGTATCTTAACTGCTGAAGTGTATTCTTTTACAAAGCAAGTTACGA

SEQ1701        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
SEQ1702        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
SEQ1703        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
SEQ1704        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
SEQ1705        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
SEQ1706        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
SEQ1707        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
SEQ1708        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
SEQ1709        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
SEQ1710        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG
SEQ1711        GTCAAGTTTCTATTGTGAGTCGCTTTTTATCTAACGATGATACTGTACTCATCATTGATG

SEQ1701        ACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATT-ATTGGTCAAGCTGGA
SEQ1702        ACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATT-ATTGGTCAAGCTGGA
SEQ1703        ACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATT-ATTGGTCAAGCTGGA
SEQ1704        ACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATT-ATTGGTCAAGCTGAA
SEQ1705        ACTTTTTAACAAACGGTCAAGC--------------------------------------
SEQ1706        ACTTTTTAGCAAACMGTCYAGCGGCTAAAGGATTACTTGAAATT-ATTGGTCAAGCTGGA
SEQ1707        ACTTTTTAGCAAACGGKCAAGCGGSTAAAGGATTACTTGAAATT-ATTGGTCAAGCTGGA
SEQ1708        ACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATT-ATTGGTCAAGCTGAA
SEQ1709        ACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATTTATTGGTCAAGCTGGA
SEQ1710        ACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATT-ATTGGTCAAGCTGGA
SEQ1711        ACTTTTTAGCAAACGGTCAAGCGGCTAAAGGATTACTTGAAATT-ATTGGTCAAGCTGGA

SEQ1701        CTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTG
SEQ1702        CTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTG
SEQ1703        CTAAGGTTGCTGGTATCGGAATCYTTATTGAAAAATCTTTCCAAGATGGGCGTGATT--
SEQ1704        CTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTG
SEQ1705        -----------------------------------------------------------
SEQ1706        CTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTG
SEQ1707        CTA--------------------------------------------------------
SEQ1708        CTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTG
SEQ1709        CTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTG
SEQ1710        CTAAGGTTGCTGGTATCGGAATCGTTATTGAAAAATCTTTCCAAGATGGGCGTGATTTG
SEQ1711        CTAAGGTTGCTGGTATCGGA-------TABCMARATVSTNCSR--ATNGTSAGXANTHN

SEQ1701        TAGAAAAAACAGGTGTTCCAGT----------------
SEQ1702        TAGAAAAAACA---------------------------
SEQ1703        --------------------------------------
SEQ1704        TAGAAAAAACAGGTGTTCCGGTTACTTCTCTTGCTCGT
SEQ1705        --------------------------------------
SEQ1706        TAGAAAA-------------------------------
SEQ1707        --------------------------------------
SEQ1708        TAGAAAAAACAGGTGTTCCGGTTAC-------------
SEQ1709        TAGAAAAAACAGGTGTTCCAGT----------------
SEQ1710        TAGAAAAAACAGGTGTTCCAG-----------------
SEQ1711        HRBSYTRANSRAS-------------------------
```

Table 18: Comparative Sequences relating to SAG1600 (glutamate racemase)

SEQ ID NO. 1801: SAG1600 FROM THE H36b GBS TYPE Ib STRAIN (REVERSE COMPLEMENT)
AATCTTCATTGGAGATCAGGCTAGAGCTCCGTATGGTCCTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAATTTCT
TATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGCCTGGCAAGAAATTAAAGAAAAACTAGACGTG
CCTGTTTTAGGCGTTATTTTACCAGGAGCTAGCGCAGCTATCAAATCAACTAATTCAGGGAAAGTTGGTATTATAGGTACTCCCATGAC
TGTTAAATCAGATGCTTATCGTCAAAAAATTCAAGCTTTGTCTCCAAATACTGCTGTGGTATCCCTTGCTTGTCCGAAATTTGTTCCAA
TTGTGGAATCAAATCAGATGTCTTCTAGTTTAGCCAAAAAGGTGGTTTATGAAACGTTGTCCCCATTAGTTGGTAAATTAGATACTTTA
ATTTTAGGTTGCACGCATTATCCCTTATTACGTCCCATCATTCAAAATGTTATGGGGGCTGAGGTTAAATTAATTGATAGTGGCGCAGA
AACCGTTCGTGATATTTCTGTTTTATTGAACTATTTTGAGATAAACCATAATTGGCAAAATAAACACGGTGGTCATCACTTTTACACAA
CCGCCAGCCCAA

SEQ ID NO. 1802: SAG1600 FROM THE M732 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
AAATGTTCCGTCAACTTCCAGAAGAGGAAGTAATCTTCATTGGAGATCAGGCTAGAGCTCCGTATGGTCCTAGACCTGCTCAACAGATT
AGAGAGTTTACCTGGCAGATGGTTAACTTCTTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGC
CTGGCAAGAAATTAAAGAAAAACTAGACATCCCTGTTTTAGGCGTTATTTTACCAGGAGCTAGCGCAGCTATCAAATCAACTAATTTAG
GGAAAGTTGGTATTATAGGTACTCCCATGACTGTTAAATCAGATGCTTATCGTCAAAAAATTCAAGCTTTGTCTCCAAATACTGCTGTG
GTATCCCTTGCTTGTCCGAAATTTGTTCCAATTGTGGAATCAAATCAGATGTCTTCTAGTTTAGCCAAAAAGGTGGTTTATGAAACGTT
GTCCCCATTAGTTGGTAAATTAGATACTTTAATTTTAGGTTGCACGCATTATCCCCTATTACGTCCCATCATTCAAAATGTTATGGGGG
CTGAGGTTAAATTAATTGATAGTGGCGCAGAAACCGTTCGTGATATTTCTGTTTTATTGAACTATTTTGAGATAAACCATAATTGGCAA
AATAAACACGGTGGTCATCACTTTTACACAACCGCCAGCCCAAAAGGTTTTAAAGAAA

SEQ ID NO. 1803: SAG1600 FROM THE 090 GBS TYPE Ia STRAIN
AATCTTCATTGGAGACCAGGCTAGAGCTCCGTATGGTCCTAGACCTGCTCAACAGATTAGAGAGTtACCTGGCAGATGGTTAATTTCTT
ATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGCCTGGCAAGAAATTAAAGAAAAACTAGACATAC
CTGTTTTAGGCGTTATTTTACCAGGAGCTAGCGCAGCTATCAAATCAACTAATTCAGGGAAAGTTGGTATTATAGGTACTCCCATGACT
GTTAAATCAGATGCTTATCGTCAAAAAATTCAAGCTTTGTCTCCAAATACTGCTGTGGTATCCCTTGCTTGTCCGAAATTTGTTCCAAT
TGTGGAATCAAATCAGATGTCTTCTAGTTTAGCCAAAAAGGTGGTTTATGAAACGCTGTCCCCATTAGTTGGTAAATTAGATACTTTAA
TTTTAGGTTGCACGCATTATCCCTTATTACGTCCCATCATTCAAAATGTTATGGGGGCTGAGGTTAAATTAATTGATAGTGGCGCAGAA
ACCGTTCGTGATATTTCTGTTTTATTGAACTATTTTGAGATaAmCCATaATTGGsmAAATAAACACGGTGGTCATCACTTTTACACAAC
CGsCAGCCCAAAAGGTTTTTAAGGAAATTGCAGAACAATGGCTTAATCAAGAAATAAAT

SEQ ID NO. 1804: SAG1600 FROM THE A909 GBS TYPE Ia STRAIN
GCGGTTGTGTAAAAGTGATGACCACCGTGTTTATTTTGCCAATTATGGTTTATCTCAAAATAGTTCAATAAAACAGAAATATCACGAAC
GGTTTCTGCGCCACTATCAATTAATTTAACCTCAGCCCCCATAACATTTTGAATGATGGGACGTAATAGGGGATAATGCGTGCAACCTA
AAATTAAAGTATCTAATTTACCAACTAATGGGGACAACGTTTCATAAACCACCTTTTTGGCTAAACTAGAAGACATCTGATTTGATTCC
ACAATTGGAACAAATTTCGGACAAGCAAGGGATACCACAGCAGTATTTGGAGACAAAGCTTGAATTTTTTGACGATAAGCATCTGATTT
AACAGTCATGGGAGTACCTATAATACCAACTTTCCCTAAATTAGTTGATTTGATAGCTGCGCTAGCTCCTGGTAAAATAACGCCTAAAA
CAGGGATGTCTAGTTTTTCTTTAATTTCTTGCCAGGCAACTGCAGTTGCTGTATTACAAGCTATAACAATCATCTTAACATTTTTAGTC
AATAAGAAGTTAACCATCTGCCAGGTAAACTCTCTAATCTGTTGAGCAGGTCTAGGACCATACGGAGCTCTAGCCTGATCTCCAATGAA
GATTACTTCCTCTTCTGGAAGTTGACGGAACATTTCCTTAACAACCGTTAAACCACCT

SEQ ID NO. 1805: SAG1600 FROM THE COH1 GBS TYPE Ia STRAIN
TTCCGTCAACTTCCAAAATATGAAGTAATCTTCATTGGAGATCAGGCTAGAGCTCCGTATGGTCCTAGACCTGCTCAACAGATTAGAGA
GTTTACCTGGCAGATGGTTAACTTCTTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGCCTGGC
AAGAAATTAAAGAAAAACTAGACATCCCTGTTTTAGGCGTTATTTTACCAGGAGCTAGCGCAGCTATCAAATCAACTAATTTAGGGAAA
GTTGGTATTATAGGTACTCCCATGACTGTTAAATCAGATGCTTATCGTCAAAAAATTCAAGCTTTGTCTCCAAATACTGCTGTGGTATC
CCTTGCTTGTCCGAAAT

SEQ ID NO. 1806: SAG1600 FROM THE CJB110 GBS NONTYPEABLE STRAIN
GTAATCTTCATTGGAGATCAGGCTAGAGCTCCGTATGGTCCTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAATTT
CTTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGCCTGGCAAGAAATTAAAGAAAAACTAGACA
TAC

SEQ ID NO. 1807: SAG1600 FROM THE 1169NT1 GBS TYPE V STRAIN
CTTTTGGGCTGGCGGTTGTGTAAAATTGATGACCACCGTGTTTATTTTGCCAATTATGGTTTATCTCAAAATAGTTCAATAAAACAGAA
ATATCACGAACGGTTTCTGCGCCACTATCAATTAATTTAACCTCAGCCCCCATAACATTTTGAATAATGGGACGTAATAGGGGATAATG
CGTGCAACCTAAAATTAAAGTATCTAATTTACCAACTAATGGGGACAATGTTTCATAAACCACCTTTTTGGCTAAACTAGAAGACATCT
GATTTGATTCCACAATTGGAACAAATTTCGGACAAGCAAGGGATACCACAGCAGTATTTGGAGACAAAGCTTGAATTTTTTGACGATAA
GCATCTGATTTAACAGTCATGGGAGTACCTATAA

SEQ ID NO. 1808: SAG1600 FROM THE 1169NT1 GBS TYPE V STRAIN
GTAATCTTCATTGGGGATCAGGCTAGAGCTCCGTATGGTCCTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAATTT
CTTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTT

SEQ ID NO. 1809: SAG1600 FROM THE 18RS21 GBS TYPE II STRAIN
GAAATGTTCCGTCAACTTCCAGAAGAGGAAGTAATCTTCATTGGAGATCAGGCTAGAGCTCCGTATGGTCCTAGACCTGCTCAACAGAT
TAGAGAGTTTACCTGGCAGATGGTTAACTTCTTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTG
CCTGGCAAGAAATTAAAGAAAAACTAGACATCCCTGTTTTAGGCGTTATTTTACCAGGAGCTAGCGCAGCTATCAAATCAACTAATTTA
GGGAAAGTTGGTATTATAGGTACTCCCATGACTGTTAAATCAGATGCTTATCGTCAAAAAATTCAAGC

Table 18: Comparative Sequences relating to SAG1600 (glutamate racemase)

SEQ ID NO. 1810: SAG1600 FROM THE 18RS21 TYPE II STRAIN
ATTTCTTTAAAACCTTTTGGGCTGGCGGTTGTGTAATATTGATGACCACCGTGTTTATTTTGCCAATTATGGTTTATCTCAAAATAGTTCAATAAAACAGAAATATCACGAACGGTTTCTGCGCCACTATCAATTAATTTAACCTCAGCCCCCATAACATTTTGAATGATGGGACGTAATATGGGATAATGCGTGCAACCTAAAATTAAAGTA

SEQ ID NO. 1811: SAG1600 FROM THE 2603 V/R GBS TYPE V STRAIN
ATTTCTTTAAAACCTTTTGGGCTGGCGGTTGTGTAATAAGTGATGACCACCGTGTTTATTTTGCCAATTATGGTTTATCTCAAAATAGTTCAATAAAACAGAAATATCACGAACGGTTTCTGCGCCACTATCAATTAATTTAACCTCAGCCCCCATAACATTTTGAATGATGGGACGTAATAGGGGATAATGCGTGCAACCTAAAATTAAAGTATCTAATTTACCAACTAATGGGGACAACGTTTCATAAACCACCTTTTTGGCTAAACTAGAAGACATCTGATTTGATTCCACAATTGGAACAA

SEQ ID NO. 1812: SAG1600 FROM THE M781 GBS TYPE III STRAIN
GGCGGTTGTGTAAAAGTGATGACCACCGTGTTTATTTTGCCAATTATGGTTTATCTCAAAATAGTTCAATAAAACAGAAATATCACGAACGGTTTCTGCGCCACTATCAATTAATTTAACCTCAGCCCCCATAACATTTTGAATGATGGGACGTAATAGGGGATAATGCGTGCAACCTAAAATTAAAGTATCTAATTTACCAACTAATGGGGACAACGTTTCATAAACCACCTTTTTGGCTAAACTAGAAGA

SEQ ID NO. 1813: SAG1600 FROM THE M 781 GBS TYPE III STRAIN
AATCTTCATTGGAGATCAGGCTAGAGCTCCGTATGGTCCTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAACTTCTTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGC

SEQ ID NO. 1814: SAG1600 FROM THE JM9130013 GS TYPE VIII STRAIN
TGGGCTGGCGGTTGTGTAAAAGTGATGACCACCGTGTTTATTTTGCCAATTATGGTTTATCTCAAAATAGTTCAATAAAACAGAAATATCACGAACGGTTTCTGCGCCACTATCAATTAATTTAACCTCAGCCCCCATAACATTTTGAATGATGGGACGTAATAAGGGATAATGCGTGCAACCTAAAATTAAAGTATCTAATTTACCAACTAATGGGGACAACGTTTCATAAACCACCTTTTTGGCTAAACTAGAAGACATCTGATTTGATTCCACAATTGGAACAAATTTCGGACAAGCAAGGGATACCACAGCAGTATTTGGAGACAAAGCTTGAATTTTTTGACGATAAGCATCTGATTTAACAGTCATGGGAGTACCTATAATACCAACTTTCCCTGAA

```
SEQ1801           -------------------------------AATCTTCATTGGAGATCAGGCTAGAGCT
SEQ1802           AAATGTTCCGTCAACTTCCAGAAGAGGAAGTAATCTTCATTGGAGATCAGGCTAGAGCT
SEQ1803           -------------------------------AATCTTCATTGGAGACCAGGCTAGAGCT
SEQ1804           ------------------------------------------GCGGTTGTGTAAAAG-T
SEQ1805           -----TTCCGTCAACTTCCAAAATATGAAGTAATCTTCATTGGAGATCAGGCTAGAGCT
SEQ1806           -----------------------------GTAATCTTCATTGGAGATCAGGCTAGAGCT
SEQ1807           -------------------------------CTTTTGGGCTGGCGGTTGTGTAAAAT-T
SEQ1808           -----------------------------GTAATCTTCATTGGGGATCAGGCTAGAGCT
SEQ1809           AAATGTTCCGTCAACTTCCAGAAGAGGAAGTAATCTTCATTGGAGATCAGGCTAGAGCT
SEQ1810           ------------------ATTTCTTTAAAACCTTTTGGGCTGGCGGTTGTGTAATAT-T
SEQ1811           ------------------ATTTCTTTAAAACCTTTTGGGCTGGCGGTTGTGTAATAAGT
SEQ1812           -----------------------------------------GGCGGTTGTGTAAAAG-T
SEQ1813           -------------------------------AATCTTCATTGGAGATCAGGCTAGAGCT
SEQ1814           -----------------------------------TGGGCTGGCGGTTGTGTAAAAG-T

SEQ1801           CGTATGGTC-CTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAATTT
SEQ1802           CGTATGGTC-CTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAACTT
SEQ1803           CGTATGGTC-CTAGACCTGCTCAACAGATTAGAGAGTT-ACCTGGCAGATGGTTAATTT
SEQ1804           --GATGACCACCGTGTTTATTTTGCCAATTATGG--TTTATCTCA-AAATAGTTCA---
SEQ1805           CGTATGGTC-CTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAACTT
SEQ1806           CGTATGGTC-CTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAATTT
SEQ1807           --GATGACCACCGTGTTTATTTTGCCAATTATGG--TTTATCTCA-AAATAGTTCA---
SEQ1808           CGTATGGTC-CTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAATTT
SEQ1809           CGTATGGTC-CTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAACTT
SEQ1810           --GATGACCACCGTGTTTATTTTGCCAATTATGG--TTTATCTCA-AAATAGTTCA---
SEQ1811           --GATGACCACCGTGTTTATTTTGCCAATTATGG--TTTATCTCA-AAATAGTTCA---
SEQ1812           --GATGACCACCGTGTTTATTTTGCCAATTATGG--TTTATCTCA-AAATAGTTCA---
SEQ1813           CGTATGGTC-CTAGACCTGCTCAACAGATTAGAGAGTTTACCTGGCAGATGGTTAACTT
SEQ1814           --GATGACCACCGTGTTTATTTTGCCAATTATGG--TTTATCTCA-AAATAGTTCA---
```

Table 18: Comparative Sequences relating to SAG1600 (glutamate racemase)

```
SEQ1801          TTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGC
SEQ1802          TTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGC
SEQ1803          TTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGC
SEQ1804          --ATAAAACAGAAATATCACGAACGGT-TTCTGCGCCACTATCAATTAATTTAACCTCA
SEQ1805          TTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGC
SEQ1806          TTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGC
SEQ1807          --ATAAAACAGAAATATCACGAACGGT-TTCTGCGCCACTATCAATTAATTTAACCTCA
SEQ1808          TTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTT--
SEQ1809          TTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGCAGTTGC
SEQ1810          --ATAAAACAGAAATATCACGAACGGT-TTCTGCGCCACTATCAATTAATTTAACCTCA
SEQ1811          --ATAAAACAGAAATATCACGAACGGT-TTCTGCGCCACTATCAATTAATTTAACCTCA
SEQ1812          --ATAAAACAGAAATATCACGAACGGT-TTCTGCGCCACTATCAATTAATTTAACCTCA
SEQ1813          TTATTGACTAAAAATGTTAAGATGATTGTTATAGCTTGTAATACAGCAACTGC------
SEQ1814          --ATAAAACAGAAATATCACGAACGGT-TTCTGCGCCACTATCAATTAATTTAACCTCA

SEQ1801          TGGCAAGAAATTAAAGAAAAACTAGACGTGCCTGTTTTAGGCGTTATTTTACCAGGAGC
SEQ1802          TGGCAAGAAATTAAAGAAAAACTAGACATCCCTGTTTTAGGCGTTATTTTACCAGGAGC
SEQ1803          TGGCAAGAAATTAAAGAAAAACTAGACATACCTGTTTTAGGCGTTATTTTACCAGGAGC
SEQ1804          CCCCCATAACATTTTGAATGATGGGACGTAATAGGGGATAATGC-GTGCAACCTAAAAT
SEQ1805          TGGCAAGAAATTAAAGAAAAACTAGACATCCCTGTTTTAGGCGTTATTTTACCAGGAGC
SEQ1806          TGGCAAGAAATTAAAGAAAAACTAGACATAC----------------------------
SEQ1807          CCCCCATAACATTTTGAATAATGGGACGTAATAGGGGATAATGC-GTGCAACCTAAAAT
SEQ1808          -----------------------------------------------------------
SEQ1809          TGGCAAGAAATTAAAGAAAAACTAGACATCCCTGTTTTAGGCGTTATTTTACCAGGAGC
SEQ1810          CCCCCATAACATTTTGAATGATGGGACGTAATATGGGATAATGC-GTGCAACCTAAAAT
SEQ1811          CCCCCATAACATTTTGAATGATGGGACGTAATAGGGGATAATGC-GTGCAACCTAAAAT
SEQ1812          CCCCCATAACATTTTGAATGATGGGACGTAATAGGGGATAATGC-GTGCAACCTAAAAT
SEQ1813          -----------------------------------------------------------
SEQ1814          CCCCCATAACATTTTGAATGATGGGACGTAATAAGGGATAATGC-GTGCAACCTAAAAT

SEQ1801          AGCGCAGCTATCAAATCAACTAATTCAGGGAAAGTTGGTATTATAGGTACTCCCATGAC
SEQ1802          AGCGCAGCTATCAAATCAACTAATTTAGGGAAAGTTGGTATTATAGGTACTCCCATGAC
SEQ1803          AGCGCAGCTATCAAATCAACTAATTCAGGGAAAGTTGGTATTATAGGTACTCCCATGAC
SEQ1804          AAAGTATCTAATTTACCAACTAATGGGGACAACGTTTCATAAACCACCTTTTTGGCTAA
SEQ1805          AGCGCAGCTATCAAATCAACTAATTTAGGGAAAGTTGGTATTATAGGTACTCCCATGAC
SEQ1806          -----------------------------------------------------------
SEQ1807          AAAGTATCTAATTTACCAACTAATGGGGACAATGTTTCATAAACCACCTTTTTGGCTAA
SEQ1808          -----------------------------------------------------------
SEQ1809          AGCGCAGCTATCAAATCAACTAATTTAGGGAAAGTTGGTATTATAGGTACTCCCATGAC
SEQ1810          AAAGTA-----------------------------------------------------
SEQ1811          AAAGTATCTAATTTACCAACTAATGGGGACAACGTTTCATAAACCACCTTTTTGGCTAA
SEQ1812          AAAGTATCTAATTTACCAACTAATGGGGACAACGTTTCATAAACCACCTTTTTGGCTAA
SEQ1813          -----------------------------------------------------------
SEQ1814          AAAGTATCTAATTTACCAACTAATGGGGACAACGTTTCATAAACCACCTTTTTGGCTAA

SEQ1801          GTTAAATCAGATGCTTATCGTCAAAAAATTCAAGCTTTGTCTCCAAATACTGCTGTGGT
SEQ1802          GTTAAATCAGATGCTTATCGTCAAAAAATTCAAGCTTTGTCTCCAAATACTGCTGTGGT
SEQ1803          GTTAAATCAGATGCTTATCGTCAAAAAATTCAAGCTTTGTCTCCAAATACTGCTGTGGT
SEQ1804          CTAGAAGACATCTGATTTGATTCCACAATTGGAACAAATTTCGGACAAGCAAGGGATAC
SEQ1805          GTTAAATCAGATGCTTATCGTCAAAAAATTCAAGCTTTGTCTCCAAATACTGCTGTGGT
SEQ1806          -----------------------------------------------------------
SEQ1807          CTAGAAGACATCTGATTTGATTCCACAATTGGAACAAATTTCGGACAAGCAAGGGATAC
SEQ1808          -----------------------------------------------------------
SEQ1809          GTTAAATCAGATGCTTATCGTCAAAAAATTCAAGC------------------------
SEQ1810          -----------------------------------------------------------
SEQ1811          CTAGAAGACATCTGATTTGATTCCACAATTGGAACAA----------------------
SEQ1812          CTAGAAGA---------------------------------------------------
SEQ1813          -----------------------------------------------------------
SEQ1814          CTAGAAGACATCTGATTTGATTCCACAATTGGAACAAATTTCGGACAAGCAAGGGATAC
```

Table 18: Comparative Sequences relating to SAG1600 (glutamate racemase)

```
SEQ1801          TCCCTTGCTTGTCCGAAATTTGTTCCAATTGTGGAATCAAATCAGATGTCTTCTAGTTT
SEQ1802          TCCCTTGCTTGTCCGAAATTTGTTCCAATTGTGGAATCAAATCAGATGTCTTCTAGTTT
SEQ1803          TCCCTTGCTTGTCCGAAATTTGTTCCAATTGTGGAATCAAATCAGATGTCTTCTAGTTT
SEQ1804          ACAGCAGTATTTGGAGACAAAGCTTGAATTTTTTGACGATAAGCATCTGATTTAACAGT
SEQ1805          TCCCTTGCTTGTCCGAAAT----------------------------------------
SEQ1806          -----------------------------------------------------------
SEQ1807          ACAGCAGTATTTGGAGACAAAGCTTGAATTTTTTGACGATAAGCATCTGATTTAACAGT
SEQ1808          -----------------------------------------------------------
SEQ1809          -----------------------------------------------------------
SEQ1810          -----------------------------------------------------------
SEQ1811          -----------------------------------------------------------
SEQ1812          -----------------------------------------------------------
SEQ1813          -----------------------------------------------------------
SEQ1814          ACAGCAGTATTTGGAGACAAAGCTTGAATTTTTTGACGATAAGCATCTGATTTAACAGT

SEQ1801          GCCAAAAAGGTGGTTTATGAAACGTTGTCCCCATTAGTTGGTAAATTAGATACTTTAAT
SEQ1802          GCCAAAAAGGTGGTTTATGAAACGTTGTCCCCATTAGTTGGTAAATTAGATACTTTAAT
SEQ1803          GCCAAAAAGGTGGTTTATGAAACGCTGTCCCCATTAGTTGGTAAATTAGATACTTTAAT
SEQ1804          ATGGGAGTACCTATAATACCAACTTTCCCTAAATTAGTTGATTTGATAGCTGCGCTAGC
SEQ1805          -----------------------------------------------------------
SEQ1806          -----------------------------------------------------------
SEQ1807          ATGGGAGTACCTATAA-------------------------------------------
SEQ1808          -----------------------------------------------------------
SEQ1809          -----------------------------------------------------------
SEQ1810          -----------------------------------------------------------
SEQ1811          -----------------------------------------------------------
SEQ1812          -----------------------------------------------------------
SEQ1813          -----------------------------------------------------------
SEQ1814          ATGGGAGTACCTATAATACCAACTTTCCCTGAATABCMARATVSTNCSRATNGTSAGGT

SEQ1801          TTAGGTTGCACGCATTATCCCTTATTACGTCCCATCATTCAAAATGTTATGGGGGCTGA
SEQ1802          TTAGGTTGCACGCATTATCCCCTATTACGTCCCATCATTCAAAATGTTATGGGGGCTGA
SEQ1803          TTAGGTTGCACGCATTATCCCTTATTACGTCCCATCATTCAAAATGTTATGGGGGCTGA
SEQ1804          CCTGGTAAAATAACGCCTAAAACAGGGATGTCTAGTTTTTCTTTAATTTCTTGCCAGGC
SEQ1805          -----------------------------------------------------------
SEQ1806          -----------------------------------------------------------
SEQ1807          -----------------------------------------------------------
SEQ1808          -----------------------------------------------------------
SEQ1809          -----------------------------------------------------------
SEQ1810          -----------------------------------------------------------
SEQ1811          -----------------------------------------------------------
SEQ1812          -----------------------------------------------------------
SEQ1813          -----------------------------------------------------------
SEQ1814          AMATRACMAS-------------------------------------------------

SEQ1801          GTTAAATTAATTGATAGTGGCGCAGAAACCGTTCGTGATATTTCTGTTTTATTGAACTA
SEQ1802          GTTAAATTAATTGATAGTGGCGCAGAAACCGTTCGTGATATTTCTGTTTTATTGAACTA
SEQ1803          GTTAAATTAATTGATAGTGGCGCAGAAACCGTTCGTGATATTTCTGTTTTATTGAACTA
SEQ1804          ACTGCAGTTGCTGTATTACAAGCTATAACAATCATCTTAACATTTTTAGTCAATAAGAA
SEQ1805          -----------------------------------------------------------
SEQ1806          -----------------------------------------------------------
SEQ1807          -----------------------------------------------------------
SEQ1808          -----------------------------------------------------------
SEQ1809          -----------------------------------------------------------
SEQ1810          -----------------------------------------------------------
SEQ1811          -----------------------------------------------------------
SEQ1812          -----------------------------------------------------------
SEQ1813          -----------------------------------------------------------
SEQ1814          -----------------------------------------------------------
```

Table 18: Comparative Sequences relating to SAG1600 (glutamate racemase)

```
SEQ1801      TTTGAGATAAACCATAATTGGCAAAATAAACACGGTGGTCATCACTTTTACACAACCGC
SEQ1802      TTTGAGATAAACCATAATTGGCAAAATAAACACGGTGGTCATCACTTTTACACAACCGC
SEQ1803      TTTGAGATAAMCCATAATTGGSMAAATAAACACGGTGGTCATCACTTTTACACAACCGS
SEQ1804      TTAACCATCTGCCAGGTAAACTCTCTAATCTGTTGAGCAGGTCTAGGACCATACGGAGC
SEQ1805      -----------------------------------------------------------
SEQ1806      -----------------------------------------------------------
SEQ1807      -----------------------------------------------------------
SEQ1808      -----------------------------------------------------------
SEQ1809      -----------------------------------------------------------
SEQ1810      -----------------------------------------------------------
SEQ1811      -----------------------------------------------------------
SEQ1812      -----------------------------------------------------------
SEQ1813      -----------------------------------------------------------
SEQ1814      -----------------------------------------------------------

SEQ1801      AGCCCAA----------------------------------------------------
SEQ1802      AGCCCAAAAGGTTTTAAAGAAA-------------------------------------
SEQ1803      AGCCCAAAAGGTTTTTAAGGAAATTGCAGAACAATGGCTTAATCAAGAAATAAAT----
SEQ1804      CTAGCCTGATCTCCAATGAAGATTACTTCCTCTTCTGGAAGTTGACGGAACATTTCCTT
SEQ1805      -----------------------------------------------------------
SEQ1806      -----------------------------------------------------------
SEQ1807      -----------------------------------------------------------
SEQ1808      -----------------------------------------------------------
SEQ1809      -----------------------------------------------------------
SEQ1810      -----------------------------------------------------------
SEQ1811      -----------------------------------------------------------
SEQ1812      -----------------------------------------------------------
SEQ1813      -----------------------------------------------------------
SEQ1814      -----------------------------------------------------------

SEQ1801      ------------------
SEQ1802      ------------------
SEQ1803      ------------------
SEQ1804      ACAACCGTTAAACCACCT
SEQ1805      ------------------
SEQ1806      ------------------
SEQ1807      ------------------
SEQ1808      ------------------
SEQ1809      ------------------
SEQ1810      ------------------
SEQ1811      ------------------
SEQ1812      ------------------
SEQ1813      ------------------
SEQ1814      ------------------
```

Table 19: Comparative Sequences relating to SAG1680 (shikimate 5-dehydrogenase)

SEQ ID NO. 1901: SAG1680 FROM THE 2603 V/R GBS TYPE V STRAIN
ATCCCTAGACCATTATAAGCATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTCTTTAGGCTTGTAAAC
CAAGTCGACAACTACTAAATTCGGTGTTAAAATTTCTGGATCGTTAATTAAACTATAATTATCTAATGGCCTCATTCCTA
AACTAGTAGCATCAATATAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAACGACT
ACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTAATTTGAGCTGTTACGATTAAATAA
TCTAATTTCCGCAACTCCCTCCATAGCTGCTTGAACTGCAACTGCTTTACCTGAACCACCAATACCAGCTATTGTAATTA
TTTTATTTTTAGCACTGAAACCTTGAGCTGCTAAAGCTTTAAAACAACCAATGCCATCTGTCATATGGCCTACTAAACGT
CCGGTTCCACCTTGATTAACGATAGTATTTACAGCACCCACTAATTTAGCTTGAGGAGATAAATCATCTAGCAAAGGGAT
AACACTCTGTTTAAATGGCATTGAAACATTAACACCACGAATACCCAATGCCCTGACACCTCGAACAGCTTCTGTTAATT
TACCCTCTTCTACTTCAAATGTCAGATAGGCATAATTCATGTTTTTTTCTTGAAAAGAGGTATTCCACATTAACGGGGAT
AGAGAGTGGCGTGCAGG

SEQ ID NO. 1902: SAG1680 FROM THE H36b GBS TYPE Ib STRAIN
GTTATTAATTGAAATGCTTCTGCTCCTTGATAAATCAGCATCCCTAGACCATTATAAGCATGTTTCACTCCATTTTGTCT
AACAAATCGTAACAATGCTGTTtCTTTAGGCTTGTAAACCAAGTCGACAACTACTAAATTCGGTGTTAAAATTTCTGGAT
CGTTAATTAAACTATAATTATCTAATGGCCTCATTCCTAAACTAGTAGCATCAATATAAAAATGACTAGTTCTAATAGCG
TCTTTAAATGCTGTCTTATTTTCTAGATAATCAACGACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTC
AATGACCTTATCGTAATTTGAGCTGTTACGATTAAATAATCTAATTTCCGCAACTCCCTCCATAGCTGCTTGAACTGCAA
CTGCTTTACCTGAACCACCAATACCAGCTATTGTAATTATTTTATTTTTAGCACTGAAACCTTGAGCTGCTAAAGCTTTA
AAACAACCAATGCCATCTGTCATATGGCCTACTAAACGTCCGGTTCCACCTTGATTAACGATAGTATTTACAGCACCCAC
TAATTTAGCTTGAGGAGATAAATCATCTAGCAAAGGGATAACACTCTGTTTAAATGGCATTGAAACATTAACACCACGAA
TACCCAATGCCCTGACACCTCGAACAGCTTCTGTTAATTTACCCTCTTCTACTTCAAATGTCAGATAGGCATAATTCATG
TTTTTTTCTTGAAAAGAGGTATTCCACATTAACGGGGATAGAGAGTGGCGTGCAGGA

SEQ ID NO. 1903: SAG1680 FROM THE M732 GBS TYPE III STRAIN
CTGGTCTAATTGCCAATCCTGCACGCCACTCTCTATCCCCGTTAATGTGGAATACCTCTTTTCAAGAAAAAAACATGAAT
TATGCCTATCTGACATTTGAAGTAGAAGAGGGTAAATTAACAGAAGCTGTTCGAGGTGTCAGGGCATTGAGTATTCGTGG
TGTTAATGTTTCAATGCCATTTAAACAGAGTGTTATCCCTTTGCTAGATGATTTATCTCCTCAAGCTAAATTAGTGGGTG
CTGTAAATACTATCGTTAATCAAGGTGGAACCGGACGTTTAGTAGGCCATATGACAGATGGCATTGGTTGTTTTAAAGCT
TTAGCAGCTCAAGGTTTCAGTGCTAAAAATAAAATAATTACAATAGCTGGTATTGGTGGTTCAGGTAAAGCAGTTGCAGT
TCAAGCAGCTATGGAGGGAGTTGCGGAAATTAGATTATTTAATCGTAACAGCTCAAATTACGATAAGGTCATTGACTTAT
CAGATAAAATTAAAAAACAGTTTCAAATAAAGGTAGTCGTTGATTATCTAGAAAATAAGACAGCATTTAAAGACGCTATT
AGAACTAGTCATTTTTATATTGATGCTACTAGTTTAGGAATGAGGCCATTAGATAATTATAGTTTAATTAACGATCCAGA
TATTTTAACACCGAATTTAGTAGTTGTCGACTT

SEQ ID NO. 1904: SAG1680 FROM THE M781 GBS TYPE III STRAIN
AAATCAGCATCCCTAGACATTATAAGCATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTCTTTAGGCT
TGTAAACCAAGTCGACAACTACTAAATTCGGTGTTAAAATTTCTGGATCGTTAATTAAACTATAATTATCTAATGGCCTC
ATTCCTAAACTAGTAGCATCAATATAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATC
AACGACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTAATTTGAGCTGTTACGAT
TAAATAATCTAATTTCCGCAACTCCCTCCATAGCTGCTTGAACTGCAACTGCTTTACCTGAACCACCAATACCAGCTATT
GTAATTATTTTATTTTTAGCACTGAAACCTTGAGCTGCTAAAGCTTTAAAACAACCAATGCCATCTGTCATATGGCCTAC
TAAACGTCCGGTTCCACCTTGATTAACGATAGTATTTACAGCACCCACTAATTTAGCTTGAGGAGATAAATCATCTAGCA
AAGGGATAACACTCTGTTTAAATGGCATTGAAACATTAACACCACGAATACTCAATGCCCTGACACCTCGAACAGCTTCT
GTTAATTTACCCTCTTCTACTTCAAATGTCAGATAGGCATAATTCATGTTTTTTTCTTGAAAAGAGGTATTCCACATTAA
CGGGGATAGAGAGTGGCGTGCA

SEQ ID NO. 1905: SAG1680 FROM THE 090 GBS TYPE Ia STRAIN
GTTCGAGGTGTCAGGGCATTGGGTATTCGTGGTGTTAATGTTTCAATGCCATTTAAACAGAGTGTTATCCCtTTGCTArA
TGATTTATCTCCTCAAGCTAAATTAGTGGGTGCTGTAAATACTATCGTTAATCAAGGTGGAACCGsACGTTTAGTAGGCC
ATATGACAGATGGCATTGGTTGTTTTAAAGCTTTAGCAGCTCAAGGTTTCAGTGCTAAAAATAAAATAGTTACAATAGCT
GGTATTGGTGGTTCAGGTAAAGCAGTTGCAGTTCAAGCAGCTATGGAGGGAGTTGCGGAAATTAGATTATTTAATCGTAA
TAGCTCAAATTACGATAAGGTCATTGACTTATCAGATAAAATTAAAAAACAGTTTCAAATAAAGGTAGTCGTTGATTATC
TAGAAAATAAGACAGCATTTAAAGACGCTATTAGAACTAGTCATTTTTATATTGATGCTACTAGTTTAGGAATGArGCCA
TTAGATAATTATAGTTTAATTAACGATCCAGAAATTTTAACACCCAATTTAGTAGTTGTCGACTTGGTTTACAAGCCTAA
AGAAACAGCATTGTTACGATTTGTTAGACAAAATGGAGTGAAACATGCTTATAATGGTCTAGGGATGCTGATTTATCAAG
GAGCAGA

Table 19: Comparative Sequences relating to SAG1680 (shikimate 5-dehydrogenase)

SEQ ID NO. 1906: SAG1680 FROM THE A909 GBS TYPE Ia STRAIN
CCCTAGACCATTATAATCATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTCTTTAGGCTTGTAAACCA
AGTCGACAACTACTAAATTCGGTGTTAAAATTTCTGGATCGTTAATTAAACTATAATTATCTAATGGCCTCATTCCTAAA
CTAGTAGCATCAATATAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAACGACTAC
CTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTAATTTGAGCTGTTACGATTAAATAATC
TAATTTCCGCAACTCCCTCCATAGCTGCTTGAACTGCAACTGCTTTACCTGAACCACCAATACCAGCTATTGTAATTATT
TTATTTTTAGCACTGAAACCTTGAGCTGCTAAAGCTTTAAAACAACCAATGCCATCTGTCATATGGCCTACTAAACGTCC
GGTTCCACCTTGATTAACGATAGTATTTACAGCACCCACTAATTTAGCTTGAGGAGATAAATCATCTAGCAAAGGATAA
CACTCTGTTTAAATGGCATTGAAACATTAACACCACGAATACCCAATGCCCTGACACCTCGAACAGCTTCTGTTAATTTA
CCCTCTTCTACTTCAAATGTCAGATAGGCATAATTCATGTTTTTTTCTTGAAAAGAGGTATTCCACATTAACGGGGATAG

SEQ ID NO. 1907: SAG1680 FROM THE COH1 GBS TYPE Ia STRAIN
TGCACGCCACTCTCTATCCCCGTTAATGTGGAATACCTCTTTTAAGAAAAAAACATGAATTATGCCTATCTGACATTTGA
AGTAGAAGAGGGTAAATTAACAGAAGCTGTTCGAGGTGTCAGGGCATTGAGTATTCGTGGTGTTAATGTTTCAATGCCAT
TTAAACAGAGTGTTATCCCTTTGCTAGATGATTTATCTCCTCAAGCTAAATTAGTGGGTGCTGTAAATACT

SEQ ID NO. 1908: SAG1680 FROM THE CJB110 GBS NONTYPEABLE STRAIN
ATTCGTTATTAATTGAAATGCTTCTGCTCCTTGATAAATCAGCATCCCTAGACCATTATAAGCATGTTTCACTCCATTTT
GTCTAACAAATCGTAACAATGCTGTTTCTTTAGGCTTGTAAACCAAGTCGACAACTACTAAATTGGGTGTTAAAATTTCT
GGATCGTTAATTAAACTATAATTATCTAATGGCCTCATTCCTAAACTAGTAGCATCAATATAAAAATGACTAGTTCTAAT
AGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAACGACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATA
AGTCAATGACCTTATCGTAATTTGAGCTATTACGATTAAATAATCTAATTTCCGCAACTCCCTCCATAACTGCTTGAACT
GCAACTGCTTTACCTGAACCACCAATACCAGCTATTGTAACTATTTT

SEQ ID NO. 1909: SAG1680 FROM THE CJB110 GBS NONTYPEABLE STRAIN
ACTCTCTATCCCCGTTAATGTGGAATACCTCTTTTCAAGAAAAAAACATGAATTATGCCTATCTGACATTTGAAGTAGAA
GAGGGTAAATTAACAGAAGCTGTTCGAGGTGTCAGGGCATTGGGTATTCGTGGTGTTAATGTTTCAATGCCATTTAAACA
GAGTGTTATCCCTTTGCTAGATGATTTATCTCCTCAAGCTAAATTAGTGGGTGCTGTAAATACTATCGTTAATCAAGGTG
GAACCGGACGTTTAGTAGGCCATATGACAGATGGCATTGGTTGTTTTAAAGCTTTAGCAGCTCAAGGTTTCAGTGCTAAA
AATAAAATAGTTACAATAGCTGGTATTGGTG

SEQ ID NO. 1910: SAG1680 FROM THE 1169NT1 GBS TYPE V STRAIN
ATTCGTTATTAATTGAAATGCTTCTGCTCCTTGATAAATCAGCATCCCTAGACCATTATAAGCATGTTTCACTCCATTTT
GTCTAACAAATCGTAACAATGCTGTTTCTTTAGGCTTGTAAACCAAGTCGACAACTACTAAATTCGGTGTTAAAATTTCT
GGATCGTTAATTAAACTATAATTATCTAATGGCCTCATTCCTAAACTAGTAGCATCAATATAAAAATGACTAGTTCTAAT
AGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAACGACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATA
AGTCAATGACCTTATCGTAATTTGAGCTGTTACGAT

SEQ ID NO. 1911: SAG1680 FROM THE 1169NT1 GBS TYPE V STRAIN
ACTTCTCTATTCCCCGTTAATGTGGAATACCTCTTTTCAAGAAAAAAACATGAATTATGCCTATCTGACATTTGAAGTAG
AAGAGGGTAAATTAACAGAAGCTGTTCGAGGTGTCAGGGCATTGGGTATTCGTGGTGTTAATGTTTCAATGCCATTTAAA
CAGAGTGTTATCCCTTTGCTAGATGATTTATCTCCTCAAGCTAAATTAGTGGGTGCTGTAAATACTATCGTTAATCAAGG
TGGAACC

SEQ ID NO. 1912: SAG1680 FROM THE 18RS21 GBS TYPE II STRAIN
TCGTTATTAATTGAAATGCTTCTGCTCCTTGATAAATCATCATCCCTAGACCATTATAAGCATGTTTCACTCCATTTTGT
CTAACAAATCGTAACAATGCTGTTTCTTTAGGCTTGTAAACCAAGTCGACAACTACTAAATTCGGTGTTAAAATTTCTGG
ATCGTTAATTAAACTATAATTATCTAATGGCCTCATTCCTAAACTAGTAGCATCAATATAAAAATGACTAGTTCTAATAG
CGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAACGACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAG
TCAATGACCTTATCGTAATTTGAGCTGTTACGATTAAATAATCTAATTTCCGCAAC

SEQ ID NO. 1913: SAG1680 FROM THE 18RS21 GBS TYPE II STRAIN
ATGCCTATCTGACATTTGAAGTAGAAGAGGGTAAATTAACAGAAGCTGTTCGAGGTGTCAGGGCATTGGGTATTCGTGGT
GTTAATGTTTCAATGCCATTTAAACAGAGTGTTATCCCTTTGCTAGATGATTTATCTCCTCAAGCTAAATTAGTGGGTGC
TGTAAATACTATCGTTAATCAAGGTGGAACCGGACGTTTAGTAGGCCATATGACAGATGGCATTGGTTGTTTTAAAGCTT
TAGCAGCTCAAGGTTTCAGTGCTAAAAATAAAATAATTACAATAGCTGGTATTGGTGGTTCAGGTAAAGCAGTTGCAGTT
CAAGCAGCTATGGAGGGAGTTGCGG

Table 19: Comparative Sequences relating to SAG1680 (shikimate 5-dehydrogenase)

SEQ ID NO. 1914: SAG1680 FROM THE JM9130013 GBS TYPE VIII STRAIN

```
CCCTAGACCATTATAAGTCATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTCTTTAGGCTTGTAAACC
AAGTCGACAACTACTAAATTGGGTGTTAAAATTTCTGGATCGTTAATTAAACTATAATTATCTAATGGCCTCATTCCTAA
ACTAGTAGCATCAATATAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAACGACTA
CCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTAATTTGAGCTATTACGATTAAATAAT
CTAATTTCCGCAACTCCCTCCATAGCTGCTTGAACTGCAACTGCTTTACCTGAACCACCAATACCAGCTATTGTAACTAT
TTTATTTTTAGCACTGAAACCTTGAGCTGCTAAAGCTTTAAAACAACCAATGCCATCTGTCAT
```

```
SEQ1901       -----------------------------------------------------ATCCCT
SEQ1902       --------------GTTATTAATTGAAATGCTTCTGCTCCTTGATAAATCAGCATCCCT
SEQ1903       TGGTCTAATTGCCAATCCTGCACGCCACTCTCTAT-CCCCGTTAATGTGGAATACCTCT
SEQ1904       ---------------------------------------------AAATCAGCATCCCT
SEQ1905       -----------------------------------------------------------
SEQ1906       -------------------------------------------------------CCCT
SEQ1907       ------------------TGCACGCCACTCTCTAT-CCCCGTTAATGTGGAATACCTCT
SEQ1908       ----------ATTCGTTATTAATTGAAATGCTTCTGCTCCTTGATAAATCAGCATCCCT
SEQ1909       --------------------------ACTCTCTAT-CCCCGTTAATGTGGAATACCTCT
SEQ1910       ----------ATTCGTTATTAATTGAAATGCTTCTGCTCCTTGATAAATCAGCATCCCT
SEQ1911       -------------------------ACTTCTCTATTCCCCGTTAATGTGGAATACCTCT
SEQ1912       ------------TCGTTATTAATTGAAATGCTTCTGCTCCTTGATAAATCATCATCCCT
SEQ1913       -----------------------------------------------------------
SEQ1914       -------------------------------------------------------CCCT

SEQ1901       GACCATTATAAG-CATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTC
SEQ1902       GACCATTATAAG-CATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTC
SEQ1903       TTCAAGAAAAAAACATGAATTATGCCTATCTGACATTTGAAGTAGAAGAGGGTAAATTA
SEQ1904       GAC-ATTATAAG-CATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTC
SEQ1905       -----------------------------------------------------------
SEQ1906       GACCATTATAAT-CATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTC
SEQ1907       TT-AAGAAAAAAACATGAATTATGCCTATCTGACATTTGAAGTAGAAGAGGGTAAATTA
SEQ1908       GACCATTATAAG-CATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTC
SEQ1909       TTCAAGAAAAAAACATGAATTATGCCTATCTGACATTTGAAGTAGAAGAGGGTAAATTA
SEQ1910       GACCATTATAAG-CATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTC
SEQ1911       TTCAAGAAAAAAACATGAATTATGCCTATCTGACATTTGAAGTAGAAGAGGGTAAATTA
SEQ1912       GACCATTATAAG-CATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTC
SEQ1913       ---------------------ATGCCTATCTGACATTTGAAGTAGAAGAGGGTAAATTA
SEQ1914       GACCATTATAAGTCATGTTTCACTCCATTTTGTCTAACAAATCGTAACAATGCTGTTTC

SEQ1901       TTAGGCTTGTAAACCAAGTC--GACAACTACTAAATTCGGTGTTAAAATTTCTGGATCG
SEQ1902       TTAGGCTTGTAAACCAAGTC--GACAACTACTAAATTCGGTGTTAAAATTTCTGGATCG
SEQ1903       CAGAAGCTGTTCGAGGTGTCAGGGCATTGAGTATTCGTGGTGTTAATGTTTCAATGCCA
SEQ1904       TTAGGCTTGTAAACCAAGTC--GACAACTACTAAATTCGGTGTTAAAATTTCTGGATCG
SEQ1905       --------GTTCGAGGTGTCAGGGCATTGGGTATTCGTGGTGTTAATGTTTCAATGCCA
SEQ1906       TTAGGCTTGTAAACCAAGTC--GACAACTACTAAATTCGGTGTTAAAATTTCTGGATCG
SEQ1907       CAGAAGCTGTTCGAGGTGTCAGGGCATTGAGTATTCGTGGTGTTAATGTTTCAATGCCA
SEQ1908       TTAGGCTTGTAAACCAAGTC--GACAACTACTAAATTGGGTGTTAAAATTTCTGGATCG
SEQ1909       CAGAAGCTGTTCGAGGTGTCAGGGCATTGGGTATTCGTGGTGTTAATGTTTCAATGCCA
SEQ1910       TTAGGCTTGTAAACCAAGTC--GACAACTACTAAATTCGGTGTTAAAATTTCTGGATCG
SEQ1911       CAGAAGCTGTTCGAGGTGTCAGGGCATTGGGTATTCGTGGTGTTAATGTTTCAATGCCA
SEQ1912       TTAGGCTTGTAAACCAAGTC--GACAACTACTAAATTCGGTGTTAAAATTTCTGGATCG
SEQ1913       CAGAAGCTGTTCGAGGTGTCAGGGCATTGGGTATTCGTGGTGTTAATGTTTCAATGCCA
SEQ1914       TTAGGCTTGTAAACCAAGTC--GACAACTACTAAATTGGGTGTTAAAATTTCTGGATCG
```

Table 19: Comparative Sequences relating to SAG1680 (shikimate 5-dehydrogenase)

```
SEQ1901          TT-AATTAAACTATAATTATCT-----AATGGCCTCATTCCT-AAACTAGTAGCATCAAT
SEQ1902          TT-AATTAAACTATAATTATCT-----AATGGCCTCATTCCT-AAACTAGTAGCATCAAT
SEQ1903          TTTAAACAGAGTGTTATCCCTTTGCTAGATGATTTATCTCCTCAAGCTAAATTAGTGGGT
SEQ1904          TT-AATTAAACTATAATTATCT-----AATGGCCTCATTCCT-AAACTAGTAGCATCAAT
SEQ1905          TTTAAACAGAGTGTTATCCCTTTGCTARATGATTTATCTCCTCAAGCTAAATTAGTGGGT
SEQ1906          TT-AATTAAACTATAATTATCT-----AATGGCCTCATTCCT-AAACTAGTAGCATCAAT
SEQ1907          TTTAAACAGAGTGTTATCCCTTTGCTAGATGATTTATCTCCTCAAGCTAAATTAGTGGGT
SEQ1908          TT-AATTAAACTATAATTATCT-----AATGGCCTCATTCCT-AAACTAGTAGCATCAAT
SEQ1909          TTTAAACAGAGTGTTATCCCTTTGCTAGATGATTTATCTCCTCAAGCTAAATTAGTGGGT
SEQ1910          TT-AATTAAACTATAATTATCT-----AATGGCCTCATTCCT-AAACTAGTAGCATCAAT
SEQ1911          TTTAAACAGAGTGTTATCCCTTTGCTAGATGATTTATCTCCTCAAGCTAAATTAGTGGGT
SEQ1912          TT-AATTAAACTATAATTATCT-----AATGGCCTCATTCCT-AAACTAGTAGCATCAAT
SEQ1913          TTTAAACAGAGTGTTATCCCTTTGCTAGATGATTTATCTCCTCAAGCTAAATTAGTGGGT
SEQ1914          TT-AATTAAACTATAATTATCT-----AATGGCCTCATTCCT-AAACTAGTAGCATCAAT

SEQ1901          TAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAAC
SEQ1902          TAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAAC
SEQ1903          CTGTAAATACTATCGTTAATCAAGGTGGAACCGGACGTTTAGTAGGCCATATGACAGAT
SEQ1904          TAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAAC
SEQ1905          CTGTAAATACTATCGTTAATCAAGGTGGAACCGSACGTTTAGTAGGCCATATGACAGAT
SEQ1906          TAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAAC
SEQ1907          CTGTAAATACT-------------------------------------------------
SEQ1908          TAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAAC
SEQ1909          CTGTAAATACTATCGTTAATCAAGGTGGAACCGGACGTTTAGTAGGCCATATGACAGAT
SEQ1910          TAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAAC
SEQ1911          CTGTAAATACTATCGTTAATCAAGGTGGAACC----------------------------
SEQ1912          TAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAAC
SEQ1913          CTGTAAATACTATCGTTAATCAAGGTGGAACCGGACGTTTAGTAGGCCATATGACAGAT
SEQ1914          TAAAAATGACTAGTTCTAATAGCGTCTTTAAATGCTGTCTTATTTTCTAGATAATCAAC

SEQ1901          ACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTA
SEQ1902          ACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTA
SEQ1903          GCATTGGTTGTTTTAAAGCTTTAGCAGCTCAAGGTTTCAGTGCTAAAAATAAAATAATT
SEQ1904          ACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTA
SEQ1905          GCATTGGTTGTTTTAAAGCTTTAGCAGCTCAAGGTTTCAGTGCTAAAAATAAAATAGTT
SEQ1906          ACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTA
SEQ1907          ------------------------------------------------------------
SEQ1908          ACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTA
SEQ1909          GCATTGGTTGTTTTAAAGCTTTAGCAGCTCAAGGTTTCAGTGCTAAAAATAAAATAGTT
SEQ1910          ACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTA
SEQ1911          ------------------------------------------------------------
SEQ1912          ACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTA
SEQ1913          GCATTGGTTGTTTTAAAGCTTTAGCAGCTCAAGGTTTCAGTGCTAAAAATAAAATAATT
SEQ1914          ACTACCTTTATTTGAAACTGTTTTTTAATTTTATCTGATAAGTCAATGACCTTATCGTA

SEQ1901          TTTGAGCTGTTACGATTAAATAATCTAATTTCCGCAACTCCCTCCATAGCTGCTTGAAC
SEQ1902          TTTGAGCTGTTACGATTAAATAATCTAATTTCCGCAACTCCCTCCATAGCTGCTTGAAC
SEQ1903          CAATAGCTGGTATTGGTGGTTCAGGTAAAGCAGTTGCAGTTCAAGCAGCTATGGAGGGA
SEQ1904          TTTGAGCTGTTACGATTAAATAATCTAATTTCCGCAACTCCCTCCATAGCTGCTTGAAC
SEQ1905          CAATAGCTGGTATTGGTGGTTCAGGTAAAGCAGTTGCAGTTCAAGCAGCTATGGAGGGA
SEQ1906          TTTGAGCTGTTACGATTAAATAATCTAATTTCCGCAACTCCCTCCATAGCTGCTTGAAC
SEQ1907          ------------------------------------------------------------
SEQ1908          TTTGAGCTATTACGATTAAATAATCTAATTTCCGCAACTCCCTCCATAACTGCTTGAAC
SEQ1909          CAATAGCTGGTATTGGTG------------------------------------------
SEQ1910          TTTGAGCTGTTACGAT--------------------------------------------
SEQ1911          ------------------------------------------------------------
SEQ1912          TTTGAGCTGTTACGATTAAATAATCTAATTTCCGCAAC----------------------
SEQ1913          CAATAGCTGGTATTGGTGGTTCAGGTAAAGCAGTTGCAGTTCAAGCAGCTATGGAGGGA
SEQ1914          TTTGAGCTATTACGATTAAATAATCTAATTTCCGCAACTCCCTCCATAGCTGCTTGAAC
```

Table 19: Comparative Sequences relating to SAG1680 (shikimate 5-dehydrogenase)

```
SEQ1901        GCAACTGCTTTACCTGAACCACCAATACCAGCTATTGTAATTATTTTATTTTTAGCACT
SEQ1902        GCAACTGCTTTACCTGAACCACCAATACCAGCTATTGTAATTATTTTATTTTTAGCACT
SEQ1903        TTGCGGAAATTAGATTATTTAATCGTAACAGCTCAAATTACGATAAGGTCATTGACTTA
SEQ1904        GCAACTGCTTTACCTGAACCACCAATACCAGCTATTGTAATTATTTTATTTTTAGCACT
SEQ1905        TTGCGGAAATTAGATTATTTAATCGTAATAGCTCAAATTACGATAAGGTCATTGACTTA
SEQ1906        GCAACTGCTTTACCTGAACCACCAATACCAGCTATTGTAATTATTTTATTTTTAGCACT
SEQ1907        ------------------------------------------------------------
SEQ1908        GCAACTGCTTTACCTGAACCACCAATACCAGCTATTGTAACTATTTT-------------
SEQ1909        ------------------------------------------------------------
SEQ1910        ------------------------------------------------------------
SEQ1911        ------------------------------------------------------------
SEQ1912        ------------------------------------------------------------
SEQ1913        TTGCGG------------------------------------------------------
SEQ1914        GCAACTGCTTTACCTGAACCACCAATACCAGCTATTGTAACTATTTTATTTTTAGCACT

SEQ1901        AAACCTTGAGCTGCTAAAGCTTTAAAACAACCAATGCCATCTGTCATATGGCCTACTAA
SEQ1902        AAACCTTGAGCTGCTAAAGCTTTAAAACAACCAATGCCATCTGTCATATGGCCTACTAA
SEQ1903        CAGATAAAATTAAAAAACAGTTTCAAATAAAGGTAGTCGTTGATTATCTAGAAAATAAG
SEQ1904        AAACCTTGAGCTGCTAAAGCTTTAAAACAACCAATGCCATCTGTCATATGGCCTACTAA
SEQ1905        CAGATAAAATTAAAAAACAGTTTCAAATAAAGGTAGTCGTTGATTATCTAGAAAATAAG
SEQ1906        AAACCTTGAGCTGCTAAAGCTTTAAAACAACCAATGCCATCTGTCATATGGCCTACTAA
SEQ1907        ------------------------------------------------------------
SEQ1908        ------------------------------------------------------------
SEQ1909        ------------------------------------------------------------
SEQ1910        ------------------------------------------------------------
SEQ1911        ------------------------------------------------------------
SEQ1912        ------------------------------------------------------------
SEQ1913        ------------------------------------------------------------
SEQ1914        AAACCTTGAGCTGCTAAAGCTTTAAAACAACCAATGCCATCTGTCAT-TABCMARAT--

SEQ1901        CGTCCGGTTCCACCTTGATTAACGATAGTATTTACAGCACCCACTAATTTAGCTTGAGG
SEQ1902        CGTCCGGTTCCACCTTGATTAACGATAGTATTTACAGCACCCACTAATTTAGCTTGAGG
SEQ1903        CAGCATTTAAAGACGCTATTAGAACTAGTCATTTTTATATTGATGCTACTAGTTTAGGA
SEQ1904        CGTCCGGTTCCACCTTGATTAACGATAGTATTTACAGCACCCACTAATTTAGCTTGAGG
SEQ1905        CAGCATTTAAAGACGCTATTAGAACTAGTCATTTTTATATTGATGCTACTAGTTTAGGA
SEQ1906        CGTCCGGTTCCACCTTGATTAACGATAGTATTTACAGCACCCACTAATTTAGCTTGAGG
SEQ1907        ------------------------------------------------------------
SEQ1908        ------------------------------------------------------------
SEQ1909        ------------------------------------------------------------
SEQ1910        ------------------------------------------------------------
SEQ1911        ------------------------------------------------------------
SEQ1912        ------------------------------------------------------------
SEQ1913        ------------------------------------------------------------
SEQ1914        STNCSRATNGTSASHKMATDHYDRGNAS--------------------------------

SEQ1901        GATAAATCATCTAGCAAAGGGATAACACTCTGTTTAAATGGCATTGAAACATTAACACC
SEQ1902        GATAAATCATCTAGCAAAGGGATAACACTCTGTTTAAATGGCATTGAAACATTAACACC
SEQ1903        TGAGGCCATTAGATAATTATAGTTTAATTAACGATCCAGATATTTTAACACCGAATTTA
SEQ1904        GATAAATCATCTAGCAAAGGGATAACACTCTGTTTAAATGGCATTGAAACATTAACACC
SEQ1905        TGARGCCATTAGATAATTATAGTTTAATTAACGATCCAGAAATTTTAACACCCAATTTA
SEQ1906        GATAAATCATCTAGCAAAGGGATAACACTCTGTTTAAATGGCATTGAAACATTAACACC
SEQ1907        ------------------------------------------------------------
SEQ1908        ------------------------------------------------------------
SEQ1909        ------------------------------------------------------------
SEQ1910        ------------------------------------------------------------
SEQ1911        ------------------------------------------------------------
SEQ1912        ------------------------------------------------------------
SEQ1913        ------------------------------------------------------------
SEQ1914        ------------------------------------------------------------
```

Table 19: Comparative Sequences relating to SAG1680 (shikimate 5-dehydrogenase)

```
SEQ1901          CGAATACCCAATGCCCTGACACCTCGAACAGCTTCTGTTAATTTACCCTCTTCTACTTC
SEQ1902          CGAATACCCAATGCCCTGACACCTCGAACAGCTTCTGTTAATTTACCCTCTTCTACTTC
SEQ1903          TAGTTGTCGACTT-----------------------------------------------
SEQ1904          CGAATACTCAATGCCCTGACACCTCGAACAGCTTCTGTTAATTTACCCTCTTCTACTTC
SEQ1905          TAGTTGTCGACTTGGTTTACAAGCCTAAAGAAACAGCATTGTTACGATTTGTTAGACAA
SEQ1906          CGAATACCCAATGCCCTGACACCTCGAACAGCTTCTGTTAATTTACCCTCTTCTACTTC
SEQ1907          ------------------------------------------------------------
SEQ1908          ------------------------------------------------------------
SEQ1909          ------------------------------------------------------------
SEQ1910          ------------------------------------------------------------
SEQ1911          ------------------------------------------------------------
SEQ1912          ------------------------------------------------------------
SEQ1913          ------------------------------------------------------------
SEQ1914          ------------------------------------------------------------

SEQ1901          AATGTCAGATAGGCATAATTCATGTTTTTTTCTTGAAAAGAGGTATTCCACATTAACGG
SEQ1902          AATGTCAGATAGGCATAATTCATGTTTTTTTCTTGAAAAGAGGTATTCCACATTAACGG
SEQ1903          ------------------------------------------------------------
SEQ1904          AATGTCAGATAGGCATAATTCATGTTTTTTTCTTGAAAAGAGGTATTCCACATTAACGG
SEQ1905          ATGGAGTGAAACATGCTTATAATGGTCTAGGGATGCTGATTTATCAAGGAGCAGA----
SEQ1906          AATGTCAGATAGGCATAATTCATGTTTTTTTCTTGAAAAGAGGTATTCCACATTAACGG
SEQ1907          ------------------------------------------------------------
SEQ1908          ------------------------------------------------------------
SEQ1909          ------------------------------------------------------------
SEQ1910          ------------------------------------------------------------
SEQ1911          ------------------------------------------------------------
SEQ1912          ------------------------------------------------------------
SEQ1913          ------------------------------------------------------------
SEQ1914          ------------------------------------------------------------

SEQ1901          GATAGAGAGTGGCGTGCAGG-
SEQ1902          GATAGAGAGTGGCGTGCAGGA
SEQ1903          ---------------------
SEQ1904          GATAGAGAGTGGCGTGCA---
SEQ1905          ---------------------
SEQ1906          GATAG----------------
SEQ1907          ---------------------
SEQ1908          ---------------------
SEQ1909          ---------------------
SEQ1910          ---------------------
SEQ1911          ---------------------
SEQ1912          ---------------------
SEQ1913          ---------------------
SEQ1914          ---------------------
```

Table 20: Comparative Sequences relating to SAG1723 (signal peptidase I)

SEQ ID NO. 2001: SAG1723 FROM THE COH1 GBS TYPE Ia STRAIN
ATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCGGCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTC
ATCAAATATAAAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATACTAAATTATTTAAAAAGGA
TAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAAGACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCA
GCGAATTTACTACTGTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGATAGTCGTGCCGTCGGTTCC
TTCAAAA

SEQ ID NO. 2002: SAG1680 FROM THE CJB110 GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)
TAAAGTTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAGTTCTCAAACAAACAAAAATCAATCGATTCGATA
TTGTAGTGGCTAACGAAGAAGAAGGCGGCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATAAA
AATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATACTAAATTATTTAAAAAGGATAAATTACAGGA
AAAATATTCGTATAACCCACTTTTCCAAGACCTAGCACAAAGCTCTACCGCTTTCACTACTGACAGCAATGGCAGCAGCGAATTTACTA
CTGTCGTGCCTAAAGGCCACTATTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGATAGTCGTGCCGTCGGTCCCTTCAAAAAATCA
ACAATTGTGGGAG

SEQ ID NO. 2003: SAG1680 FROM THE 18RS21 GBS TYPE II STRAIN
TTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAGTTCTCAAACAAACAAAAATCAATCGATTCGATATTGTA
GTGGCTAACGAAGAAGAAGGCGGCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATAAAAATGA
CACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATACTAAATTATTTAAAAAGGATAAATTACAGGAAAAAT
ATTCGTATAACCCACTTTTCCAAGACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACTACTGTC
GTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGATAGTCGTGCCGTCGGTCCCTTCAAAAAATCAACGAT
TGTGGGAGAGGT

SEQ ID NO. 2004: SAG1680 FROM THE 2603 V/R GBS TYPE V STRAIN (REVERSE COMPLEMENT)
AAGTTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAGTTCTCAAACAAACAAAAATCAATCGATTCGATATT
GTAGTGGCTAACGAAGAAGAAGGCGGCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATAAAAA
TGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATACTAAATTATTTAAAAAGGATAAATTACAGGAAA
AATATTCGTATAACCCACTTTTCCAAGACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACTACT
GTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGATAGTCGTGCCGTCGGT

SEQ ID NO. 2005: SAG1680 FROM THE M732 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
TTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAGTTCTCAAACAAACAAAATAATCGATTCGATATTGTAGT
GGCTAACGAAGAAGAAGGCGGCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATAAAAATGACA
CCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATACTAAATTATTTAAAAAGGATAAATTACAGGAAAAATAT
TCGTATAACCCACTTTTCCAAGACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACTACTGTCGT
GCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGA

SEQ ID NO. 2006: SAG1680 FROM THE M781 GBS TYPE III STRAIN
TTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAGTTCTCAAACAAACAAAAATCAATCGATTCGATATTGTA
GTGGCTAACGAAGAAGAAGGCGGCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATAAAAATGA
CACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATACTAAATTATTTAAAAAGGATAAATTACAGGAAAAA
TATTCGTATAACCCACTTTTCCAAGACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACT

SEQ ID NO. 2007: SAG1680 FROM THE 1169NT1 GBS TYPE V STRAIN (REVERSE COMPLEMENT)
TTGGTAAAGTTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAGTTCTCAAACAAACAAAAATCAATCGATTC
GATATTGTAGTGGCTAACGAAGAAGAAGGCGGCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATA
TAAAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATACTAAATTATTTAAAAAGGATAAATTAC
AGGAAAAATATTCGTATAACCCACTTTTCCAAGACCTAGCACAAAGCTCTACCGCTTTCACTACTGACAGCAATGGCAGCAGCGAATTT
ACCACTGTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGATAGTCGTGCCGTCGGCCCCTTCAAAAA
ATCAACG

SEQ ID NO. 2008: SAG1680 FROM THE H36b GBS TYPE Ib STRAIN (REVERSE COMPLEMENT)
TTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAGTTCTCAAACAAACAAAAATCAATCGATTCGATATTGTA
GTGGCTAACGAAGAAGAAGGCGGCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATAAAAATGA
CACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATACTAAATTATTTAAAAAGGATAAATTACAGGAAAAAT
ATTCGTATAACCCACTTTTCCAAGACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACTACTGTC
GTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGA

SEQ ID NO. 2009: SAG1680 FROM THE 090 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
TAAAGTTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAGTTCTCAAACAAACAAAAATCAATCGATTCGATA
TTGTAGTGGCTAACGAAGAAGAAGGCGGCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATAAA
AATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATACTAAATTATTTAAAAAGGATAAATTACAGGA
AAAATATTCGTATAACCCACTTTTCCAAGACCTAGCACAAAGCTCTACCGCTTTCACTACTGACAGCAATGGCAGCAGCGAATTTACTA
CTGTCGTGCCTAAAGGCCACTATTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGATAGTCGTGCCGTCGGT

Table 20: Comparative Sequences relating to SAG1723 (signal peptidase I)

```
SEQ ID NO. 2010:  SAG1680 FROM THE A909 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
AAAGTTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAGTTCTCAAACAAACAAAAATCAATCGATTCGATAT
TGTAGTGGCTAACGAAGAAGAAGGCGGCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATAAAA
ATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATACTAAATTATTTAAAAAGGATAAATTACAGGAA
AAATATTCGTATAACCCACTTTTCCAAGACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACTAC
TGTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGATAGTCGTGCCGTCGGTCCCTTCAAAAAATCAA
CG

SEQ2001          -----------------------------------------------------------
SEQ2002          ---TAAAGTTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAG
SEQ2003          --------TTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAG
SEQ2004          -----AAGTTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAG
SEQ2005          --------TTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAG
SEQ2006          --------TTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAG
SEQ2007          TGGTAAAGTTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAG
SEQ2008          --------TTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAG
SEQ2009          ---TAAAGTTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAG
SEQ2010          ----AAAGTTGACGGACACTCCATGGATCCAACTTTAGCTGACAAGGAACAGCTAGTAG

SEQ2001          --------------------ATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCG
SEQ2002          TCTCAAACAAACAAAAATCAATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCG
SEQ2003          TCTCAAACAAACAAAAATCAATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCG
SEQ2004          TCTCAAACAAACAAAAATCAATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCG
SEQ2005          TCTCAAACAAACAAAA--TAATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCG
SEQ2006          TCTCAAACAAACAAAAATCAATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCG
SEQ2007          TCTCAAACAAACAAAAATCAATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCG
SEQ2008          TCTCAAACAAACAAAAATCAATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCG
SEQ2009          TCTCAAACAAACAAAAATCAATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCG
SEQ2010          TCTCAAACAAACAAAAATCAATCGATTCGATATTGTAGTGGCTAACGAAGAAGAAGGCG

SEQ2001          GCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATA
SEQ2002          GCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATA
SEQ2003          GCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATA
SEQ2004          GCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATA
SEQ2005          GCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATA
SEQ2006          GCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATA
SEQ2007          GCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATA
SEQ2008          GCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATA
SEQ2009          GCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATA
SEQ2010          GCCAAAAGAAAAAAATTGTTAAACGTGTCATTGGTATGCCAGGTGATGTCATCAAATATA

SEQ2001          AAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATA
SEQ2002          AAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATA
SEQ2003          AAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATA
SEQ2004          AAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATA
SEQ2005          AAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATA
SEQ2006          AAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATA
SEQ2007          AAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATA
SEQ2008          AAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATA
SEQ2009          AAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATA
SEQ2010          AAAATGACACCTTAACTATTAACAATAAAAAAACAGAAGAACCTTACCTCAAGGAATATA

SEQ2001          CTAAATTATTT-AAAAAGGATAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAA
SEQ2002          CTAAATTATTT-AAAAAGGATAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAA
SEQ2003          CTAAATTATTT-AAAAAGGATAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAA
SEQ2004          CTAAATTATTT-AAAAAGGATAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAA
SEQ2005          CTAAATTATTT-AAAAAGGATAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAA
SEQ2006          CTAAATTATTTTAAAAAGGATAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAA
SEQ2007          CTAAATTATTT-AAAAAGGATAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAA
SEQ2008          CTAAATTATTT-AAAAAGGATAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAA
SEQ2009          CTAAATTATTT-AAAAAGGATAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAA
SEQ2010          CTAAATTATTT-AAAAAGGATAAATTACAGGAAAAATATTCGTATAACCCACTTTTCCAA
```

Table 20: Comparative Sequences relating to SAG1723 (signal peptidase I)

```
SEQ2001        GACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACT
SEQ2002        GACCTAGCACAAAGCTCTACCGCTTTCACTACTGACAGCAATGGCAGCAGCGAATTTACT
SEQ2003        GACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACT
SEQ2004        GACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACT
SEQ2005        GACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACT
SEQ2006        GACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACT
SEQ2007        GACCTAGCACAAAGCTCTACCGCTTTCACTACTGACAGCAATGGCAGCAGCGAATTTACC
SEQ2008        GACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACT
SEQ2009        GACCTAGCACAAAGCTCTACCGCTTTCACTACTGACAGCAATGGCAGCAGCGAATTTACT
SEQ2010        GACCTAGCACAAAGCTCTACCGCTTTCACCACTGACAGCAATGGCAGCAGCGAATTTACT

SEQ2001        CTGTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGAT
SEQ2002        CTGTCGTGCCTAAAGGCCACTATTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGAT
SEQ2003        CTGTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGAT
SEQ2004        CTGTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGAT
SEQ2005        CTGTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGA---------------
SEQ2006        -----------------------------------------------------------
SEQ2007        CTGTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGAT
SEQ2008        CTGTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGA---------------
SEQ2009        CTGTCGTGCCTAAAGGCCACTATTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGAT
SEQ2010        CTGTCGTGCCTAAAGGCCACTACTATCTTGTTGGTGATGACCGAATTGTCTCTAAAGAT

SEQ2001        GTCGTGCCGTCGGTTCCTTCAAAA-----------------------------------
SEQ2002        GTCGTGCCGTCGGTCCCTTCAAAAAATCAACAATTGTGGGAG-----------------
SEQ2003        GTCGTGCCGTCGGTCCCTTCAAAAAATCAACGATTGTGGGAGAGGT-------------
SEQ2004        GTCGTGCCGTCGGT---------------------------------------------
SEQ2005        -----------------------------------------------------------
SEQ2006        -----------------------------------------------------------
SEQ2007        GTCGTGCCGTCGGCCCCTTCAAAAAATCAACG---------------------------
SEQ2008        -----------------------------------------------------------
SEQ2009        GTCGTGCCGTCGGT---------------------------------------------
SEQ2010        GTCGTGCCGTCGGTCCCTTCAAAAAATCAACGTABCMARATVSTNCSRATNGTSAGSGN

SEQ2001        ----
SEQ2002        ----
SEQ2003        ----
SEQ2004        ----
SEQ2005        ----
SEQ2006        ----
SEQ2007        ----
SEQ2008        ----
SEQ2009        ----
SEQ2010        TDAS
```

Table 21: Comparative Sequences relating to SAG0079 (adenylate kinase)

SEQ ID NO. 2101: SAG0079 FROM THE 2603V/R GBS TYPE V STRAIN
AATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTTGAAGAATTTGGTGTTGCTCACATCTCA
ACAGGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTT
CCTGATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCA
CGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGAT
CCATCATGTCTTATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTA
GATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGA
GAACCTATTCTTGAACACTATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTTGCAGATGTT
GAAAAAGCGTTG

SEQ ID NO. 2102: SAG0079 FROM THE 090 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
AATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTTGAAGAATTTGGTGTTGCTCACATCTCA
ACAGGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTT
CCTGATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCA
CGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGAT
CCATCATGTCTTATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTA
GATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGA
GAACCTATTCTTGAACACTATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTTGCAGATGTT
GAAAAAGCGTTGCTAGAACTCAAA

SEQ ID NO. 2103: SAG0079 FROM THE 1169NT1 GBS TYPE V STRAIN (REVERSE COMPLEMENT)
TGGTAAAGGGACTCAAGCAGCTAAGATTGTTGAAGAATTTGGTGTTGCGCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGC
TAATCAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGATCAAGTAACAAACGGGATTGTAAA
AGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGGTATCCACGTACTATTGAACAAGCACACGCCTTAGA
TGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCATCATGTCTTATAGAGCGTTTGAGTGG
TCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACG
TGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTCATATTGCTCAAGGAGAACCTATTCTTGAACACTATAGTAAGCT
TGGCCTTGTTACAGATATTGAAGGTAATCAAGAAATAA

SEQ ID NO. 2104: SAG0079 FROM THE 18RS21 GBS TYPE II STRAIN (REVERSE COMPLEMENT)
AATCTTTTAACCACGGGTTCGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTTGAAGAATTTGGTGTTGCTCACATCTCA
ACAGGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTT
CCTGATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCA
CGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGAT
CCATCATGTCTTATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTA
GATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGA
GAACCTATTCTTGAACACTATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTTGCAGATGTT
GAAAAAGCGTTGCTAGAA

SEQ ID NO. 2105: SAG0079 FROM THE 2603V/R GBS TYPE V STRAIN (REVERSE COMPLEMENT)
AATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTTGAAGAATTTGGTGTTGCTCACATCTCA
ACAGGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTT
CCTGATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCA
CGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGAT
CCATCATGTCTTATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTA
GATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGA
GAACCTATTCTTGAACACTATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTTGCAGATGTT
GAAAAAGCGTTG

SEQ ID NO. 2106: SAG0079 FROM THE A909 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
AATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTTGAAGAATTTGGTGTTGCTCACATCTCA
ACAGGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTT
CCTGATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCA
CGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGAT
CCATCATGTCTTATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTA
GATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGA
GAATCTATTCTTGAACACTATCGAAAGCTTGGTCTTGTTACAGATATTGAAGGTAA

SEQ ID NO. 2107: SAG0079 FROM THE CJB110 GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)
AATCTTTTAACCACGGGTTTGCTTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTTGAAGAATTTGGTGTTGCTCACATCTCA
ACAGGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTT
CCTGATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCA
CGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGAT
CCATCATGTCTTATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTA
GATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGA
GAACCTATTCTTGAACACTATAG

Table 21: Comparative Sequences relating to SAG0079 (adenylate kinase)

SEQ ID NO. 2108: SAG0079 FROM THE COH1 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
ATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATTGTTGAAGAATTTGGTGTTGCTCACATCTCAA
CAGGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCCAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTC
CTGATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCAC
GTACTATTGAGCAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATC
CAACATGCCTTATAGAGCGTTTGAGTGGCCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTAG
ATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAG
AACCTATTCTTGAACACTATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTTGCAGATGTTG
AAAAAGCGTTGCTAG

SEQ ID NO. 2109: SAG0079 FROM THE H36b GBS TRYP Ib STRAIN (REVERSE COMPLEMENT)
CAGGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTC
CTGATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCAC
GTACTATTGAACAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATC
CATCATGTCTTATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTAG
ATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAG
AATCTATTCTTGAACACTATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTTGCAGATGTTG
AAAAAGCGTTGCT

SEQ ID NO. 2110: SAG0079 FROM THE JM9130013 GBS TYPE VIII STRAIN (REVERSE COMPLEMENT)
AATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTTGAAGAATTTGGTGTTGCTCACATCTCA
ACAGGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTT
CCTGATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCA
CGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGAT
CCATCATGTCTTATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTA
GATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTTAAACGTCGCTTGGACGTTAATATTGCTCAAGGA
GAACCTATTCTTGAACACTATAAAAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCA

SEQ ID NO. 2111: SAG0079 FROM THE M732 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
CTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATTGTTGAAGAATTTGGTGTTGCTCACATCTCAACA
GGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCCAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCT
GATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCACGT
ACTATTGAGCAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCA
ACATGCCTTATAGAGCGTTTGAGTGGCCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTAGAT
TATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAA
CCTATTCTTGAACACTATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTTGCAGATGTTGAA
AAAGCGTTGCTAGAACTCAAA

SEQ ID NO. 2112: SAG0079 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
AATCTTTTAATTACGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATTGTTGAAGAATTTGGTGTTGCTCACATCTCA
ACAGGGGATATGTTCCGCGCCGCAATGGCTAATCAAACCCAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTT
CCTGATGAAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGTTTTTTACTTGATGGATATCCA
CGTACTATTGAGCAAGCACACGCCTTAGATGCTACGCTTGAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGAT
CCAACATGCCTTATAGAGCGTTTGAGTGGCCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTGTTCAACCCACCAGTA
GATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCTGAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAA

```
SEQ2101        ATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTT
SEQ2102        ATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTT
SEQ2103        ----------------------------TGGTAAAGGGACTCAAGCAGCTAAGATTGTT
SEQ2104        ATCTTTTAACCACGGGTTCGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTT
SEQ2105        ATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTT
SEQ2106        ATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTT
SEQ2107        ATCTTTTAACCACGGGTTTGCTTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTT
SEQ2108        ATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATTGTT
SEQ2109        -----------------------------------------------------------
SEQ2110        ATCTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATCGTT
SEQ2111        --CTTTTAATTATGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATTGTT
SEQ2112        ATCTTTTAATTACGGGTTTGCCTGGTGCTGGTAAAGGTACTCAAGCAGCTAAGATTGTT

SEQ2101        AAGAATTTGGTGTTGCTCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT
SEQ2102        AAGAATTTGGTGTTGCTCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT
SEQ2103        AAGAATTTGGTGTTGCGCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT
SEQ2104        AAGAATTTGGTGTTGCTCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT
SEQ2105        AAGAATTTGGTGTTGCTCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT
SEQ2106        AAGAATTTGGTGTTGCTCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT
SEQ2107        AAGAATTTGGTGTTGCTCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT
SEQ2108        AAGAATTTGGTGTTGCTCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT
```

Table 21: Comparative Sequences relating to SAG0079 (adenylate kinase)

```
SEQ2109          ---------------------------CAGGGGATATGTTCCGCGCCGCAATGGCTAAT
SEQ2110          AAGAATTTGGTGTTGCTCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT

SEQ2111          AAGAATTTGGTGTTGCTCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT
SEQ2112          AAGAATTTGGTGTTGCTCACATCTCAACAGGGGATATGTTCCGCGCCGCAATGGCTAAT

SEQ2101          CAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2102          CAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2103          CAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2104          CAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2105          CAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2106          CAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2107          CAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2108          CAAACCCAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2109          CAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2110          CAAACCGAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2111          CAAACCCAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT
SEQ2112          CAAACCCAAATGGGACGTTTAGCTAAAAGTTATATTGATAAAGGTGAATTGGTTCCTGAT

SEQ2101          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2102          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2103          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2104          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2105          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2106          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2107          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2108          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2109          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2110          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2111          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT
SEQ2112          AAGTAACAAACGGGATTGTAAAAGAGCGCTTAGCTGAGGATGATATCGCAGAAAAAGGT

SEQ2101          TTTTTACTTGATGGATATCCACGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTT
SEQ2102          TTTTTACTTGATGGATATCCACGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTT
SEQ2103          TTTTTACTTGATGGGTATCCACGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTT
SEQ2104          TTTTTACTTGATGGATATCCACGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTT
SEQ2105          TTTTTACTTGATGGATATCCACGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTT
SEQ2106          TTTTTACTTGATGGATATCCACGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTT
SEQ2107          TTTTTACTTGATGGATATCCACGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTT
SEQ2108          TTTTTACTTGATGGATATCCACGTACTATTGAGCAAGCACACGCCTTAGATGCTACGCTT
SEQ2109          TTTTTACTTGATGGATATCCACGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTT
SEQ2110          TTTTTACTTGATGGATATCCACGTACTATTGAACAAGCACACGCCTTAGATGCTACGCTT
SEQ2111          TTTTTACTTGATGGATATCCACGTACTATTGAGCAAGCACACGCCTTAGATGCTACGCTT
SEQ2112          TTTTTACTTGATGGATATCCACGTACTATTGAGCAAGCACACGCCTTAGATGCTACGCTT

SEQ2101          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCATCATGTCTT
SEQ2102          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCATCATGTCTT
SEQ2103          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCATCATGTCTT
SEQ2104          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCATCATGTCTT
SEQ2105          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCATCATGTCTT
SEQ2106          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCATCATGTCTT
SEQ2107          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCATCATGTCTT
SEQ2108          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCAACATGCCTT
SEQ2109          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCATCATGTCTT
SEQ2110          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCATCATGTCTT
SEQ2111          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCAACATGCCTT
SEQ2112          GAAGAACTAGGACTACGCTTAGATGGTGTTATTAATATTAAAGTGGATCCAACATGCCTT

SEQ2101          ATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2102          ATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2103          ATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2104          ATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2105          ATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2106          ATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2107          ATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2108          ATAGAGCGTTTGAGTGGCCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2109          ATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2110          ATAGAGCGTTTGAGTGGTCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2111          ATAGAGCGTTTGAGTGGCCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
SEQ2112          ATAGAGCGTTTGAGTGGCCGTATTATCAATCGTAAAACTGGTGAAACTTTCCACAAAGTG
```

Table 21: Comparative Sequences relating to SAG0079 (adenylate kinase)

```
SEQ2101         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2102         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2103         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2104         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2105         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2106         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2107         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2108         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2109         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2110         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2111         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT
SEQ2112         TTCAACCCACCAGTAGATTATAAAGAAGAAGATTACTATCAACGTGAAGATGATAAGCCT

SEQ2101         GAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAACCTATTCTTGAACAC
SEQ2102         GAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAACCTATTCTTGAACAC
SEQ2103         GAAACTGTCAAACGTCGCTTGGACGTTCATATTGCTCAAGGAGAACCTATTCTTGAACAC
SEQ2104         GAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAACCTATTCTTGAACAC
SEQ2105         GAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAACCTATTCTTGAACAC
SEQ2106         GAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAATCTATTCTTGAACAC
SEQ2107         GAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAACCTATTCTTGAACAC
SEQ2108         GAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAACCTATTCTTGAACAC
SEQ2109         GAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAATCTATTCTTGAACAC
SEQ2110         GAAACTGTTAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAACCTATTCTTGAACAC
SEQ2111         GAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAAGGAGAACCTATTCTTGAACAC
SEQ2112         GAAACTGTCAAACGTCGCTTGGACGTTAATATTGCTCAATABCMARATVSTNCSR---AT

SEQ2101         ATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTT
SEQ2102         ATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTT
SEQ2103         ATAGTAAGCTTGGCCTTGTTACAGATATTGAAGGTAATCAAGAAATAA-----------
SEQ2104         ATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTT
SEQ2105         ATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTT
SEQ2106         ATCGAAAGCTTGGTCTTGTTACAGATATTGAAGGTAA----------------------
SEQ2107         ATAG-------------------------------------------------------
SEQ2108         ATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTT
SEQ2109         ATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTT
SEQ2110         ATAAAAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCA-------------------
SEQ2111         ATCGTAAGCTTGGTCTTGTTACAGATATTGAAGGTAATCAAGAAATAACAGAAGTTTTT
SEQ2112         GTSAGADNYATKNAS--------------------------------------------

SEQ2101         CAGATGTTGAAAAAGCGTTG------------
SEQ2102         CAGATGTTGAAAAAGCGTTGCTAGAACTCAAA
SEQ2103         --------------------------------
SEQ2104         CAGATGTTGAAAAAGCGTTGCTAGAA------
SEQ2105         CAGATGTTGAAAAAGCGTTG------------
SEQ2106         --------------------------------
SEQ2107         --------------------------------
SEQ2108         CAGATGTTGAAAAAGCGTTGCTAG--------
SEQ2109         CAGATGTTGAAAAAGCGTTGCT----------
SEQ2110         --------------------------------
SEQ2111         CAGATGTTGAAAAAGCGTTGCTAGAACTCAAA
SEQ2112         --------------------------------
```

Table 21: Comparative Sequences relating to SAG0079 (adenylate kinase)

>SEQ ID N0 2150:090 frame: 1
NLLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
YRKLGLVTDIEGNQEITEVFADVEKALLELK

>SEQ ID NO 2151:114_1169NT frame: 2
GKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPDQVTNGIVKER
LAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCLIERLSGRIIN
RKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVHIAQGEPILEHYSKLGLVTDI
EGNQEI

>SEQ ID NO 2152: 114_18RS21 frame: 1
NLLTTGSPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
YRKLGLVTDIEGNQEITEVFADVEKALLE

>SEQ ID NO 2153: 114_2603 frame: 1
NLLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
YRKLGLVTDIEGNQEITEVFADVEKAL

>SEQ ID NO 2154: 114_A909 frame: 1
NLLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGESILEH
YRKLGLVTDIEG

>SEQ ID NO 2155:114_A909 frame: 1
NLLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGESILEH
YRKLGLVTDIEG

>SEQ ID NO 2156: 114_CJB110 frame: 1
NLLTTGLLGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
Y

>SEQ ID NO 2157: 114_COH1 frame: 3
LLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTQMGRLAKSYIDKGELVPDE
VTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPTCLI
ERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEHY
RKLGLVTDIEGNQEITEVFADVEKALL

>SEQ ID NO 2158: 114_H36B frame: 3
GDMFRAAMANQTEMGRLAKSYIDKGELVPDEVTNGIVKERLAEDDIAEKGFLLDGYPRTI
EQAHALDATLEELGLRLDGVINIKVDPSCLIERLSGRIINRKTGETFHKVFNPPVDYKEE
DYYQREDDKPETVKRRLDVNIAQGESILEHYRKLGLVTDIEGNQEITEVFADVEKAL

>SEQ ID NO 2159: 114_JM9130013 frame: 1
NLLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
YKKLGLVTDIEGN

>SEQ ID NO 2160:114_M732 frame: 1
LLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTQMGRLAKSYIDKGELVPDE
VTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPTCLI
ERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEHY
RKLGLVTDIEGNQEITEVFADVEKALLELK

>SEQ ID NO 2161: 114_M781 frame: 1
NLLITGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTQMGRLAKSYIDKGELVPD
EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPTCL
IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQ

Table 21: Comparative Sequences relating to SAG0079 (adenylate kinase)

```
SEQ2150         LLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
SEQ2151         ---------GKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
SEQ2152         LLTTGSPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
SEQ2153         LLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
SEQ2154         LLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
SEQ2155         LLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
SEQ2156         LLTTGLLGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
SEQ2157         LLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTQMGRLAKSYIDKGELVPD
SEQ2158         -----------------------------GDMFRAAMANQTEMGRLAKSYIDKGELVPD
SEQ2159         LLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTEMGRLAKSYIDKGELVPD
SEQ2160         LLIMGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTQMGRLAKSYIDKGELVPD
SEQ2161         LLITGLPGAGKGTQAAKIVEEFGVAHISTGDMFRAAMANQTQMGRLAKSYIDKGELVPD

SEQ2150         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
SEQ2151         QVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
SEQ2152         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
SEQ2153         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
SEQ2154         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
SEQ2155         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
SEQ2156         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
SEQ2157         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPTCL
SEQ2158         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
SEQ2159         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPSCL
SEQ2160         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPTCL
SEQ2161         EVTNGIVKERLAEDDIAEKGFLLDGYPRTIEQAHALDATLEELGLRLDGVINIKVDPTCL

SEQ2150         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
SEQ2151         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVHIAQGEPILEH
SEQ2152         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
SEQ2153         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
SEQ2154         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGESILEH
SEQ2155         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGESILEH
SEQ2156         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
SEQ2157         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
SEQ2158         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGESILEH
SEQ2159         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
SEQ2160         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQGEPILEH
SEQ2161         IERLSGRIINRKTGETFHKVFNPPVDYKEEDYYQREDDKPETVKRRLDVNIAQ-------

SEQ2150         RKLGLVTDIEGNQEITEVFADVEKALLELK
SEQ2151         SKLGLVTDIEGNQEI---------------
SEQ2152         RKLGLVTDIEGNQEITEVFADVEKALLE--
SEQ2153         RKLGLVTDIEGNQEITEVFADVEKAL----
SEQ2154         RKLGLVTDIEG-------------------
SEQ2155         RKLGLVTDIEG-------------------
SEQ2156         ------------------------------
SEQ2157         RKLGLVTDIEGNQEITEVFADVEKALL---
SEQ2158         RKLGLVTDIEGNQEITEVFADVEKAL----
SEQ2159         KKLGLVTDIEGN------------------
SEQ2160         RKLGLVTDIEGNQEITEVFADVEKALLELK
SEQ2161         ------------------------------
```

Table 22: Comparative Sequences relating to SAG0093 (D-alanyl-D-alanine carboxypeptidase family protein)

SEQ ID NO. 2201: SAG0093 FROM THE 090 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
AAGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATATCCTCTCAAAAAAGAAATAAGAAATTA
CAATTACCAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTT
CCTGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGAA
CATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAAT
TTGACGAGGGGACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGAT
ATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTA
CGGTTTCCGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATAT
ATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAA

SEQ ID NO. 2202: SAG0093 FROM THE 1169NT1 GBS TYPE V STRAIN (REVERSE COMPLEMENT)
AAGCCTAACAGTCAACAATCATCACCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATATCCTCTCAAAAAAGAAATAAGAAATTA
CGATTACCAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTG
CCTGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGAA
CATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAAT
TTGACGAGGGGACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGAT
ATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTA
CGGTTTCCGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATAT
ATGGCCGAACATCGTTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAA

SEQ ID NO. 2203: SAG0093 FROM THE 18RS21 GBS TYPE II STRAIN
AAGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATATCCTCTCAAAAAAGAAATAAGAAATTA
CAATTACCAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTT
CCTGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGAA
CATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAAT
TTGACGAGGGGACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGAT
ATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTA
CGGTTTCCGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATAT
ATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAA

SEQ ID NO. 2204: SAG0093 FROM THE 2603V/R GBS TYPE V STRAIN
ACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATATCCTCTCAAAAAAGAAATAAGAAATTACAATTAC
CAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTTCCTGTTG
AAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGAACATTTAA
TTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGA
GGGGACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGATATGAGTA
CTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTC
CGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATATATGGCCA
AACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAAAACCCAGCTTTCTTGTACAA

SEQ ID NO. 2205: SAG0093 FROM THE A909 GBS TYPE Ia STRAIN
AAGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGACATCCTCTCAAAAAAGAAATAAGAAATTA
CGATTACCAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTG
CCTGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGAA
CATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAAATGACTAGTAACCCTAAT
TTGACGAAGGAACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGAT
ATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTA
CGGTTTCCGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATAT
ATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAA

SEQ ID NO. 2206: SAG0093 FROM THE CJB110 GBS NONTYPEABLE STRAIN
AAGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATATCCTCTCAAAAAAGAAATAAGAAATTT
ACAATTACCAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGT
TCCTGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGA
ACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAA
TTTGACGAGGGGACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGA
TATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTT
ACGGTTTCCGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATA
TATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAA

Table 22: Comparative Sequences relating to SAG0093 (D-alanyl-D-alanine carboxypeptidase family protein)

SEQ ID NO. 2207: SAG0093 FROM THE COH1 GBS TYPE III STRAIN
CCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGACATCCTCTCAAAAAAGAAATTAAGAAATTACG
ATTACCAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTGCC
TGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGAACA
TTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTT
GACGAGGGGACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGATAT
GAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACG
GTTTCCGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATATAT
GGTCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAAAACCCAGCTTTCTTGTACAA

SEQ ID NO. 2208: SAG0093 FROM THE H36b GBS TYPE Ib STRAIN
AAGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGACATCCTCTCAAAAAAGAAATAAGAAATTA
CGATTACCAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTG
CCTGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGAA
CATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAwGAAATGACTAGTAACCCTAAT
TTGACGAAGGAACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGAT
ATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTA
CGGTTTCCGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATAT
ATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAA

SEQ ID NO. 2209: SAG0093 FROM THE JM9130013 GBS TYPE VIII STRAIN
AAGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATATCCTCTCAAAAAAGAAATAAGAAATTA
CAATTACCAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTT
CCTGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGAA
CATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAAT
TTGACGAGGGGACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGAT
ATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTA
CGGTTTCCGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATAT
ATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAA

SEQ ID NO. 2210: SAG0093 FROM THE M732 GBS TYPE III STRAIN
AGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGACATCCTCTCAAAAAAGAAATAAGAAATTAC
GATTACCAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTGC
CTGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGAAC
ATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATT
TGACGAGGGGACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGATA
TGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTAC
GGTTTCCGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATATA
TGGTCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAAAACCCAGCTTTCTT

SEQ ID NO. 2211: SAG0093 FROM THE M781 GBS TYPE III STRAIN
AAGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGACATCCTCTCAAAAAAGAAATAAGAAATTA
CGATTACCAGCTGTATCATCAAAAGATTGGAACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTG
CCTGTTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTGCTAGAGCAATTGATTCACGAGAA
CATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGGAGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAAT
TTGACGAGGGGACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGACTGGATTAGCGATGGAT
ATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAGTCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTA
CGGTTTCCGGATGGTAAAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGTCTGCAAAATAT
ATGGTCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGGAGAATAACCAA

```
SEQ2201         AGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATA
SEQ2202         AGCCTAACAGTCAACAATCATCACCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATA
SEQ2203         AGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATA
SEQ2204         ------ACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATA
SEQ2205         AGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGACA
SEQ2206         AGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATA
SEQ2207         --CCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGACA
SEQ2208         AGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGACA
SEQ2209         AGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGATA
SEQ2210         AGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGACA
SEQ2211         AGCCTAACAGTCAACAATCATCATCTCAAAAGTTGAGGAATGAGGATATAAAAAAGACA
```

Table 22: Comparative Sequences relating to SAG0093 (D-alanyl-D-alanine carboxypeptidase family protein)

```
SEQ2201          TCCTCTCAAAAAAGAAAT-AAGAAATT-ACAATTACCAGCTGTATCATCAAAAGATTGGA
SEQ2202          TCCTCTCAAAAAAGAAAT-AAGAAATT-ACGATTACCAGCTGTATCATCAAAAGATTGGA
SEQ2203          TCCTCTCAAAAAAGAAAT-AAGAAATT-ACAATTACCAGCTGTATCATCAAAAGATTGGA
SEQ2204          TCCTCTCAAAAAAGAAAT-AAGAAATT-ACAATTACCAGCTGTATCATCAAAAGATTGGA
SEQ2205          TCCTCTCAAAAAAGAAAT-AAGAAATT-ACGATTACCAGCTGTATCATCAAAAGATTGGA
SEQ2206          TCCTCTCAAAAAAGAAAT-AAGAAATTTACAATTACCAGCTGTATCATCAAAAGATTGGA
SEQ2207          TCCTCTCAAAAAAGAAATTAAGAAATT-ACGATTACCAGCTGTATCATCAAAAGATTGGA
SEQ2208          TCCTCTCAAAAAAGAAAT-AAGAAATT-ACGATTACCAGCTGTATCATCAAAAGATTGGA
SEQ2209          TCCTCTCAAAAAAGAAAT-AAGAAATT-ACAATTACCAGCTGTATCATCAAAAGATTGGA
SEQ2210          TCCTCTCAAAAAAGAAAT-AAGAAATT-ACGATTACCAGCTGTATCATCAAAAGATTGGA
SEQ2211          TCCTCTCAAAAAAGAAAT-AAGAAATT-ACGATTACCAGCTGTATCATCAAAAGATTGGA

SEQ2201          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTTCCTG
SEQ2202          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTGCCTG
SEQ2203          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTTCCTG
SEQ2204          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTTCCTG
SEQ2205          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTGCCTG
SEQ2206          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTTCCTG
SEQ2207          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTGCCTG
SEQ2208          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTGCCTG
SEQ2209          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTTCCTG
SEQ2210          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTGCCTG
SEQ2211          ACTTGATTTTGGTCAATCGTGACCATAAACATGAAGAATTAAGTCCAGATGTGGTGCCTG

SEQ2201          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG
SEQ2202          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG
SEQ2203          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG
SEQ2204          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG
SEQ2205          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG
SEQ2206          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG
SEQ2207          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG
SEQ2208          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG
SEQ2209          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG
SEQ2210          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG
SEQ2211          TTGAAAATATTTATTTGGATAAACGTATTACGAAGCAAGCTACTCAGTTTTTAGAGGCTG

SEQ2201          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG
SEQ2202          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG
SEQ2203          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG
SEQ2204          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG
SEQ2205          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG
SEQ2206          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG
SEQ2207          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG
SEQ2208          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG
SEQ2209          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG
SEQ2210          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG
SEQ2211          CTAGAGCAATTGATTCACGAGAACATTTAATTTCGGGTTATCGTAGTGTTGCCTATCAGG

SEQ2201          AGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGAGGG
SEQ2202          AGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGAGGG
SEQ2203          AGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGAGGG
SEQ2204          AGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGAGGG
SEQ2205          AGAAGTTGTTCAATTCTTATGTTACTCAAGAAATGACTAGTAACCCTAATTTGACGAAGG
SEQ2206          AGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGAGGG
SEQ2207          AGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGAGGG
SEQ2208          AGAAGTTGTTCAATTCTTATGTTACTCAWGAAATGACTAGTAACCCTAATTTGACGAAGG
SEQ2209          AGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGAGGG
SEQ2210          AGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGAGGG
SEQ2211          AGAAGTTGTTCAATTCTTATGTTACTCAAGAGATGACTAGTAACCCTAATTTGACGAGGG
```

Table 22: Comparative Sequences relating to SAG0093 (D-alanyl-D-alanine carboxypeptidase family protein)

```
SEQ2201          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA
SEQ2202          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA
SEQ2203          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA
SEQ2204          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA
SEQ2205          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA
SEQ2206          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA
SEQ2207          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA
SEQ2208          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA
SEQ2209          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA
SEQ2210          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA
SEQ2211          ACAAGCAGAAAAGTTGGTAAAAACTTACTCTCAGCCTGCAGGTGCTAGTGAACACCAGA

SEQ2201          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG
SEQ2202          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG
SEQ2203          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG
SEQ2204          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG
SEQ2205          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG
SEQ2206          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG
SEQ2207          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG
SEQ2208          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG
SEQ2209          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG
SEQ2210          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG
SEQ2211          CTGGATTAGCGATGGATATGAGTACTGTAGATTCTTTGAATGAGAGCGATCCTAGAGTAG

SEQ2201          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA
SEQ2202          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA
SEQ2203          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA
SEQ2204          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA
SEQ2205          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA
SEQ2206          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA
SEQ2207          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA
SEQ2208          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA
SEQ2209          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA
SEQ2210          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA
SEQ2211          TCAGTCAGTTGAAAAAGATAGCTCCACAATATGGTTTTGTCTTACGGTTTCCGGATGGTA

SEQ2201          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT
SEQ2202          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT
SEQ2203          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT
SEQ2204          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT
SEQ2205          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT
SEQ2206          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT
SEQ2207          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT
SEQ2208          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT
SEQ2209          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT
SEQ2210          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT
SEQ2211          AAACAGCAGAAACAGGGGTAGGTTATGAAGATTGGCATTACCGCTATGTTGGGGTAGAGT

SEQ2201          CTGCAAAATATATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGG
SEQ2202          CTGCAAAATATATGGCCGAACATCGTTTAACATTAGAAGAATACATAACTTTATTAAAGG
SEQ2203          CTGCAAAATATATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGG
SEQ2204          CTGCAAAATATATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGG
SEQ2205          CTGCAAAATATATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGG
SEQ2206          CTGCAAAATATATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGG
SEQ2207          CTGCAAAATATATGGTCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGG
SEQ2208          CTGCAAAATATATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGG
SEQ2209          CTGCAAAATATATGGCCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGG
SEQ2210          CTGCAAAATATATGGTCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGG
SEQ2211          CTGCAAAATATATGGTCAAACATCATTTAACATTAGAAGAATACATAACTTTATTAAAGG
```

Table 22: Comparative Sequences relating to SAG0093 (D-alanyl-D-alanine carboxypeptidase family protein)

```
SEQ2201        AGAATAACCAA-------------------------------------------------
SEQ2202        AGAATAACCAA-------------------------------------------------
SEQ2203        AGAATAACCAA-------------------------------------------------
SEQ2204        AGAATAACCAAAACCCAGCTTTCTTGTACAA-----------------------------
SEQ2205        AGAATAACCAA-------------------------------------------------
SEQ2206        AGAATAACCAA-------------------------------------------------
SEQ2207        AGAATAACCAAAACCCAGCTTTCTTGTACAA-----------------------------
SEQ2208        AGAATAACCAA-------------------------------------------------
SEQ2209        AGAATAACCAA-------------------------------------------------
SEQ2210        AGAATAACCAAAACCCAGCTTTCTT-----------------------------------
SEQ2211        AGAATAACCAATABCMARATVSTNCSRATNGTSAGDAANYDAANNCARBXYTDASAMYRT
```

>SEQ ID NO 2250: 18_090 frame: 1
KPNSQQSSSQKLRNEDIKKISSQKRNKKLQLPAVSSKDWNLILVNRDHKHEELSPDVVPV
ENIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
QAEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
TAETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQ

>SEQ ID NO 2251: 18_1169NT frame: 1
KPNSQQSSPQKLRNEDIKKISSQKRNKKLRLPAVSSKDWNLILVNRDHKHEELSPDVVPV
ENIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
QAEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
TAETGVGYEDWHYRYVGVESAKYMAEHRLTLEEYITLLKENNQ

>SEQ ID NO 2252: 18_18RS21 frame: 1
KPNSQQSSSQKLRNEDIKKISSQKRNKKLQLPAVSSKDWNLILVNRDHKHEELSPDVVPV
ENIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
QAEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
TAETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQ

>SEQ ID NO 2253: 18_2603 frame: 3
SQQSSSQKLRNEDIKKISSQKRNKKLQLPAVSSKDWNLILVNRDHKHEELSPDVVPVENI
YLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRGQAE
KLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGKTAE
TGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQNPAFLY

>SEQ ID NO 2254: 18_A909 frame: 1
KPNSQQSSSQKLRNEDIKKTSSQKRNKKLRLPAVSSKDWNLILVNRDHKHEELSPDVVPV
ENIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTKE
QAEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
TAETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQ

>SEQ ID NO 2255:18_CJB110 frame: 1
KPNSQQSSSQKLRNEDIKKISSQKRNKKFTITSCIIKRLELDFGQS

>SEQ ID NO 2256:18_COH1 frame: 1
PNSQQSSSQKLRNEDIKKTSSQKRN

>SEQ ID NO 2257: 18_H36B frame: 1
KPNSQQSSSQKLRNEDIKKTSSQKRNKKLRLPAVSSKDWNLILVNRDHKHEELSPDVVPV
ENIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTXEMTSNPNLTKE
QAEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
TAETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQ

>SEQ ID NO 2258: 18_JM9130013 frame: 1
KPNSQQSSSQKLRNEDIKKISSQKRNKKLQLPAVSSKDWNLILVNRDHKHEELSPDVVPV
ENIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
QAEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
TAETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQ

>SEQ ID NO 2259:18_M732 frame: 3
PNSQQSSSQKLRNEDIKKTSSQKRNKKLRLPAVSSKDWNLILVNRDHKHEELSPDVVPVE
NIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRGQ
AEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGKT
AETGVGYEDWHYRYVGVESAKYMVKHHLTLEEYITLLKENNQNPAF

Table 22: Comparative Sequences relating to SAG0093 (D-alanyl-D-alanine carboxypeptidase family protein)

>SEQ ID NO 2260: 18_M781 frame: 1
KPNSQQSSSQKLRNEDIKKTSSQKRNKKLRLPAVSSKDWNLILVNRDHKHEELSPDVVPV
ENIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
QAEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
TAETGVGYEDWHYRYVGVESAKYMVKHHLTLEEYITLLKENNQ

```
SEQ2250      PNSQQSSSQKLRNEDIKKISSQKRNKKLQLPAVSSKDWNLILVNRDHKHEELSPDVVPV
SEQ2251      PNSQQSSPQKLRNEDIKKISSQKRNKKLRLPAVSSKDWNLILVNRDHKHEELSPDVVPV
SEQ2252      PNSQQSSSQKLRNEDIKKISSQKRNKKLQLPAVSSKDWNLILVNRDHKHEELSPDVVPV
SEQ2253      --SQQSSSQKLRNEDIKKISSQKRNKKLQLPAVSSKDWNLILVNRDHKHEELSPDVVPV
SEQ2254      PNSQQSSSQKLRNEDIKKTSSQKRNKKLRLPAVSSKDWNLILVNRDHKHEELSPDVVPV
SEQ2255      PNSQQSSSQKLRNEDIKKISSQKRNKKFTITSCIIKRLEL-----DFGQS---------
SEQ2256      PNSQQSSSQKLRNEDIKKTSSQKRN----------------------------------
SEQ2257      PNSQQSSSQKLRNEDIKKTSSQKRNKKLRLPAVSSKDWNLILVNRDHKHEELSPDVVPV
SEQ2258      PNSQQSSSQKLRNEDIKKISSQKRNKKLQLPAVSSKDWNLILVNRDHKHEELSPDVVPV
SEQ2259      PNSQQSSSQKLRNEDIKKTSSQKRNKKLRLPAVSSKDWNLILVNRDHKHEELSPDVVPV
SEQ2260      PNSQQSSSQKLRNEDIKKTSSQKRNKKLRLPAVSSKDWNLILVNRDHKHEELSPDVVPV

SEQ2250      NIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
SEQ2251      NIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
SEQ2252      NIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
SEQ2253      NIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
SEQ2254      NIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTKE
SEQ2255      -----------------------------------------------------------
SEQ2256      -----------------------------------------------------------
SEQ2257      NIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTXEMTSNPNLTKE
SEQ2258      NIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
SEQ2259      NIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG
SEQ2260      NIYLDKRITKQATQFLEAARAIDSREHLISGYRSVAYQEKLFNSYVTQEMTSNPNLTRG

SEQ2250      AEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
SEQ2251      AEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
SEQ2252      AEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
SEQ2253      AEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
SEQ2254      AEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
SEQ2255      -----------------------------------------------------------
SEQ2256      -----------------------------------------------------------
SEQ2257      AEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
SEQ2258      AEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
SEQ2259      AEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK
SEQ2260      AEKLVKTYSQPAGASEHQTGLAMDMSTVDSLNESDPRVVSQLKKIAPQYGFVLRFPDGK

SEQ2250      AETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQ-----------------
SEQ2251      AETGVGYEDWHYRYVGVESAKYMAEHRLTLEEYITLLKENNQ-----------------
SEQ2252      AETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQ-----------------
SEQ2253      AETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQNPAFLY-----------
SEQ2254      AETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQ-----------------
SEQ2255      -----------------------------------------------------------
SEQ2256      -----------------------------------------------------------
SEQ2257      AETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQ-----------------
SEQ2258      AETGVGYEDWHYRYVGVESAKYMAKHHLTLEEYITLLKENNQ-----------------
SEQ2259      AETGVGYEDWHYRYVGVESAKYMVKHHLTLEEYITLLKENNQNPAF-------------
SEQ2260      AETGVGYEDWHYRYVGVESAKYMVKHHLTLEEYITLLKENNQTABLECMPARATIVESE

SEQ2250      -----------------------------------------------------------
SEQ2251      -----------------------------------------------------------
SEQ2252      -----------------------------------------------------------
SEQ2253      -----------------------------------------------------------
SEQ2254      -----------------------------------------------------------
SEQ2255      -----------------------------------------------------------
SEQ2256      -----------------------------------------------------------
SEQ2257      -----------------------------------------------------------
SEQ2258      -----------------------------------------------------------
SEQ2259      -----------------------------------------------------------
SEQ2260      ENCESRELATINGTSAGDALANYLDALANINECARBXYPEPTIDASEFAMILYPRTEIN
```

Table 23: Comparative Sequences relating to SAG0163
(competence protein Cg1A)

SEQ ID NO. 2301: SAG0163 FROM THE 090 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
GGCAGTAGAAGTAAATGCTCAAGATATTTATATCATTCCCAAAGGTGATTGTTATGAACTCTATATGCGTATTGATGATGAAAGGCG
GTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTAAATTTGTGGCAGGCATGAACGTTGGAGAAAAAAG
ACGAAGTCAATTAGGTTCTTGTGACTATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTGG
TCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGTTTGATAATATAAAGCAAATGAAGGAAGT
ACTGGGTACAAGAGGGCTATATCTTTTTTCCGGCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATT
TAAAAATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAACTCCAATTGAATGAGGATATTGG
AATGACTTATGATGCTTTAATCAAACTGTCTTTACGGCATCGTCCAGATATTTTAATTATCGGAGAGATTAGAGATCAAGCGACGGC
CCGTGCTGTTATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCATGCTAAAAGTATTTCCGGAGTCTATGATAGGCT
TATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTAAAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAAT
TGACTTTGAGACAGGTAACTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAAGGACATATCAG
TAAGAAACAGGCACAAGTCGAAAAAATTATCCCTCAAGAAACAACGGAAAGTAGTCCAACTTTT

SEQ ID NO. 2302: SAG0163 FROM THE 1169NT1 GBS TYPE V STRAIN (REVERSE COMPLEMENT)
GGTGATTGTTATGAAACCTCTACTATTGCGTATTTGATGATGAAAGGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTC
TTATTAGTCACTTTAAATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACTATGAACTGTCAG
AGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTGGTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTC
ATCAGGACTTAAAATATTGGTTTGATAATATAAAGCAAATGAAGGAAGTACTGGGTACAAGAGGGCTATATCTTTTTTCCGGCCCTG
TGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAAATAAGCAAATTATCACGATTGAAGATCCGGTAG
AAATCAAGAATGACAAGATGTTACAACTCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
ATCGTCCAGATATTTTAATTATCGGAGAGATTAGAGATCAAGCGACGGCTCGTGCTGTTATTCGTGCAAGTTTAACGGGAGTGATGG
TTTTTTCTACTATTCATGCTAAAAGTATTCCCGGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATA
GTCTAAAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAAGTAACTTTAAAAAACACTCATCAG
ACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAAGGATATATCAGTAAGAAACAGGCACAAGTCGAAAAAATTATCCCTCAAG
AAACAACGGAAAGTAGTCCAACTTTT

SEQ ID NO. 2303: SAG0163 FROM THE 18RS21 GBS TYPE II STRAIN (REVERSE COMPLEMENT)
GTTCAATCATTAGCAAAGCAAGTCATTCATCAGGCAGTAGAAGTAAATGCTCAAGATATTTATATCATTCCCAAAGGTGATTGTTAT
GAACTCTATATGCGTATTGATGATGAAAGGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTAAA
TTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACTATGAACTGTCAGAGGGAAGACTGGTTTCA
TTACGACTATCGAGTGTGGGAGATTATCGTGGTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATAT
TGGTTTGATAATATAAAGCAAATGAAGGAAGTACTGGGTATAAGAGGGCTATATCTTTTTTCCGGCCCTGTGGGGAGTGGTAAAACA
ACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAAATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAG
ATGTTACAACTCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGCATCGTCCAGATATTTTA
ATTATCGGAGAGATTAGAGATCAAGCGACGGCCCGTGCTGTTATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCAT
GCTAAAAGTATTCCCGGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTAAAATTAATAGCA
TATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGTAATTTTAAAAAACACTCATCAGACAAGTGGAATAGACAA
GTGGATATCTTGGCTGAAGAAGGACATATCAGTAAGAAACAGGCACAAGTCGAAAAAATTATCCCTCAAGAAACAACGGAAAGTAGT
CCAACTTTT

SEQ ID NO. 2304: SAG0163 FROM THE 2603 V/R GBS TYPE V STRAIN (REVERSE COMPLEMENT)
GATATTTATATCATTCCCAAAGGTGATTGTTATGAACTCTATATGCGTATTGATGATGAAAGGCGGTTTATTGATGTTTTTGAGTTT
AATAGGATGGCTAGTCTTATTAGTCACTTTAAATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGT
GACTATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTGGTCAAGAATCTTTAGTTATTCGT
ATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGTTTGATAATATAAAGCAAATGAAGGAAGTACTGGGTATAAGAGGGCTATAT
CTTTTTTCCGGCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAAATAAGCAAATTATCACG
ATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAACTCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATC
AAACTGTCTTTACGGCATCGTCCAGATATTTTAATTATCGGAGAGATTAGAGATCAAGCGACGGCCCGTGCTGTTATTCGTGCAAGT
TTAACGGGAGTGATGGTTTTTTCTACTATTCATGCTAAAAGTATTCCCGGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTAT
CAAGAGTTAGAAAATAGTCTAAAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGTAATTTT
AAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAAGGACATATCAGTAAGAAACAGGCACAAGTGCGA
AAAAATTATCCCTCAAGAAACAACGGAAAGTAGTCCAACTTTT

Table 23: Comparative Sequences relating to SAG0163 (competence protein Cg1A)

SEQ ID NO. 2305: SAG0163 FROM THE A909 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
GTTCAATCATTAGCAAAGCAAGTCATTCATCAGGCAGTAGAAGTAAATGCTCAAGATATTTATATCATTCCCAAAGGTGATTGTTAT
GAACTCTATATGCGTATTGATGATGAAAGGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTAAA
TTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACTATGAACTGTCAGAGGGAAGACTGGTTTCA
TTACGACTATCGAGTGTGGGAGATTATCGTGGTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATAT
TGGTTTGATAATATAAAGCAAATGAAGGAAGTACTGGGTATAAGAGGGCTATATCTTTTTTCCGGCCCTGTGGGGAGTGGTAAAACA
ACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAAATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAG
ATGTTACAACTCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGCATCGTCCAGATATTTTA
ATTATCGGAGAGATTAGAGATCAAGCGACGGCCCGTGCTGTTATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCAT
GCTAAAAGTATTCCCGGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTAAAATTAATAGCA
TATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGTAATTTTAAAAAACACTCATCAGACAAGTGGAATAGACAA
GTGGATATCTTGGCTGAAGAAGGACATATCAGTAAGAAACAGGCACAAGTCGAAAAAATTATCCCTCAAGAAACAACGGAAAGTAGT
CCAACTTTT

SEQ ID NO. 2306: SAG0163 FROM THE CJB110 GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)
GTTCAATCATTAGCAAAGCAAGTCATTCATCAGGCAGTAGAAGTAAATGCTCAAGATATTTATATCATTCCCAAAGGTGATTGTTAT
GAACTCTATATGCGTATTGATGATGAAAGGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTAAA
TTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACTATGAACTGTCAGAGGGAAGACTGGTTTCA
TTACGACTATCGAGTGTGGGAGATTATCGTGGTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATAT
TGGTTTGATAATATAAAGCAAATGAAGGAAGTACTGGGTACAAGAGGGCTATATCTTTTTTCCGGCCCTGTGGGGAGTGGTAAAACA
ACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAAATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAG
ATGTTACAACTCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGCATCGTCCAGATATTTTA
ATTATCGGAGAGATTAGAGATCAAGCGACGGCCCGTGCTGTTATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCAT
GCTAAAAGTATTTCCGGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTAAAATTAATAGCA
TATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGTAACTTTAAAAAACACTCATCAGACAAGTGGAATAGACAA
GTGGATATCTTGGCTGAAGAAGGACATATCAGTAAGAAACAGGCACAAGTCGAAAAAATTATCCCTCAAGAAACAACGGAAAGTAGT
CCAACTTTT

SEQ ID NO. 2307: SAG0163 FROM THE COH1 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
AGGTGATTGTTATGAAATTCTATATGCGTATTGATGATGAAAGGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTA
TTAGTCACTTTAAATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACTATGAACTGTCAGAGG
GAAGACTGGTTTCATTACGACTATCAAGTGTGGGAGATTATCGTGGTCAAGAATCTTTAGTTATTCGTACTTTGTATTCAGGTCATC
AGGACTTAAAATATTGGTTTGATAATATAAAGTAAATGAAGGAAGTACTGTGTGCAAGAGGGCTATATCTTTTTTCCGGCCCTGTGG
GGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAAATAAGCAAATTATCACGATTGAAGATCCGGTAGAAA
TCAAGAATGACAAGATGTTACAACTCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGCATC
GTCCAGATATTTTAATTATCGGAGAGATTAGAGATCAAGCGACGGCCCGTGCTGTTATTCGTGCAAGTTTAACGGGAGTAATGGTTT
TTTCTACTATTCATGCTAAAAGTATTCCCGGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTC
TAAAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAAGTAACTTTAAAAAACACTCATCAGACA
AGTGGAATAGACAAGTGGATATCTTGGCTGAAGAAGGACATATCAGTAAGAAACAGGCACAAGTCGAAAAAATTATCCCTCAAGAAA
CAACGGAAAGTAGTCCAACTTTT

SEQ ID NO. 2308: SAG0163 FROM THE H36b GBS TYPE Ib STRAIN (REVERSE COMPLEMENT)
TCATTAGCAAAGCAAGTCATTCATCAGGCAGTAGAAGTAAATGCTCAAGATATTTATATCATTCCCAAAGGTGATTGTTATGAACTC
TATATGCGTATTGATGATGAAAGGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTAAATTTGTG
GCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACTATGAACTGTCAGAGGGAAGACTGGTTTCATTACGA
CTATCGAGTGTGGGAGATTATCGTGGTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGTTT
GATAATATAAAGCAAATGAAGGAAGTACTGGGTATAAGAGGGCTATATCTTTTTTCCGGCCCTGTGGGGAGTGGTAAAACAACTCTC
ATGTATCAATTAGCTTCAGAAGTATTTAAAAATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTA
CAACTCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGCATCGTCCAGATATTTTAATTATC
GGAGAGAAATAGAGATCAAGCGACGGCCCGTGCTGTTATTCGTGCAAGTTTAACGGGAGTGATGTTTTTTTCTACTATTCATGCTAA
AAGTATTCCCGGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTAAAATTAATAGCATATCA
ACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGTAATTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGA
TATCTTGGCTGAAGAAGGACATATCAGTAAGAAACAGGCACAAGTCGAAAAAATTATCCCTCAAGAAACAACGGAAAGTAGTCCAAC
TTTT

Table 23: Comparative Sequences relating to SAG0163 (competence protein Cg1A)

SEQ ID NO. 2309: SAG0163 FROM THE JM9130013 GBS TYPE VIII STRAIN (REVERSE COMPLEMENT)
GTTCAATCATTAGCAAAGCAAGTCATTCATCAGGCAGTAGAAGTAAATGCTCAAGATATTTATATCATTCCCAAAGGTGATTGTTAT
GAACTCTATATGCGTATTGATGATGAAAGGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTAAA
TTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACTATGAACTGTCAGAGGGAAGACTGGTTTCA
TTACGACTATCGAGTGTGGGAGATTATCGTGGTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATAT
TGGTTTGATAATATAAAGCAAATGAAGGAAGTACTGGGTATAAGAGGGCTATATCTTTTTTCCGGCCCTGTGGGGAGTGGTAAAACA
ACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAAATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAG
ATGTTACAACTCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGCATCGTCCAGATATTTTA
ATTATCGGAGAGATTAGAGATCAAGCGACGGCCCGTGCTGTTATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCAT
GCTAAAAGTATTCCCGGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTAAAATTAATAGCA
TATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGTAATTTTAAAAAACACTCATCAGACAAGTGGAATAGACAA
GTGGATATCTTGGCTGAAGAAGGACATATCAGTAAGAAACAGGCACAAGTCGAAAAAATTATCCCTCAAGAAACAACGGAAAGTAGT
CCAACTTTT

SEQ ID NO. 2310: SAG0163 FROM THE M732 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
TGACTTGTTATGAAACTCTATATGCGTATTTGATGATGAAAAGGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTA
TTAGTCACTTTAAATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACTATGAACTGTCAGAGG
GAAGACTGGTTTCATTACGACTATCAAGTGTGGGAGATTATCGTGGTCAAGAATCTTTAGTTATTCGTACTTTGTATTCAGGTCATC
AGGACTTAAAATATTGGTTTGATAATATAAAGTAAATGAAGGAAGTACTGTGTGCAAGAGGGCTATATCTTTTTTCCGGCCCTGTGG
GGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAAATAAGCAAATTATCACGATTGAAGATCCGGTAGAAA
TCAAGAATGACAAGATGTTACAACTCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGCATC
GTCCAGATATTTTAATTATCGGAGAGATTAGAGATCAAGCGACGGCCCGTGCTGTTATTCGTGCAAGTTTAACGGGAGTAATGGTTT
TTTCTACTATTCATGCTAAAAGTATTCCCGGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTC
TAAAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAAGTAACTTTAAAAAACACTCATCAGACA
AGTGGAATAGACAAGTGGATATCTTGGCTGAAGAAGGACATATCAGTAAGAAACAGGCACAAGTCGAAAAAATTATCCCTCAAGAAA
CAACGGAAAGTAGTCCAACTTTT

SEQ ID NO. 2311: SAG0163 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
CAGTAGAAGTAAATGCTCAAGATATTTATATCATTCCCAAAGGTGATTGTTATGAATTCTATATGCGTATTGATGATGAAAGGCGGT
TTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTAAATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGAC
GAAGTCAATTAGGTTCTTGTGACTATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCAAGTGTGGGAGATTATCGTGGTC
AAGAATCTTTAGTTATTCGTACTTTGTATTCAGGTCATCAGGACTTAAAATATTGGTTTGATAATATAAAGCAAATGAAGGAAGTAC
TGTGTGCAAGAGGGCTATATCTTTTTTCCGGCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTA
AAAATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAACTCCAATTGAATGAGGATATTGGAA
TGACTTATGATGCTTTAATCAAACTGTCTTTACGGCATCGTCCAGATATTTTAATTATCGGAGAGATTAGAGATCAAGCGACGGCCC
GTGCTGTTATTCGTGCAAGTTTAACGGGAGTAATGGTTTTTTCTACTATTCATGCTAAAAGTATTCCCGGAGTCTATGATAGGCTTA
TAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTAAAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTG
ACTTTGAGACAAGTAACTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAAGGACATATCAGTA
AGAAACAGGCACAAGTCGAAAAAATTATCCCTCAAGAAACAACGGAAAGTAGTCCAACTTTT

Table 23: Comparative Sequences relating to SAG0163 (competence protein Cg1A)

```
SEQ2301     -------------------------------GGCAGTAGAAGTAAATGCTCAAGATATT
SEQ2302     -----------------------------------------------------------
SEQ2303     TTCAATCATTAGCAAAGCAAGTCATTCATCAGGCAGTAGAAGTAAATGCTCAAGATATT
SEQ2304     -----------------------------------------------------GATATT
SEQ2305     TTCAATCATTAGCAAAGCAAGTCATTCATCAGGCAGTAGAAGTAAATGCTCAAGATATT
SEQ2306     TTCAATCATTAGCAAAGCAAGTCATTCATCAGGCAGTAGAAGTAAATGCTCAAGATATT
SEQ2307     -----------------------------------------------------------
SEQ2308     -----TCATTAGCAAAGCAAGTCATTCATCAGGCAGTAGAAGTAAATGCTCAAGATATT
SEQ2309     TTCAATCATTAGCAAAGCAAGTCATTCATCAGGCAGTAGAAGTAAATGCTCAAGATATT
SEQ2310     -----------------------------------------------------------
SEQ2311     ---------------------------------CAGTAGAAGTAAATGCTCAAGATATT

SEQ2301     ATATCATTCCCAAAGGTGA-TTGTTATGAA-CTCTATA----TGCGTATT-GATGATGA
SEQ2302     --------------GGTGA-TTGTTATGAA-ACCTCTACTATTGCGTATTTGATGATGA
SEQ2303     ATATCATTCCCAAAGGTGA-TTGTTATGAA-CTCTATA----TGCGTATT-GATGATGA
SEQ2304     ATATCATTCCCAAAGGTGA-TTGTTATGAA-CTCTATA----TGCGTATT-GATGATGA
SEQ2305     ATATCATTCCCAAAGGTGA-TTGTTATGAA-CTCTATA----TGCGTATT-GATGATGA
SEQ2306     ATATCATTCCCAAAGGTGA-TTGTTATGAA-CTCTATA----TGCGTATT-GATGATGA
SEQ2307     -------------AGGTGA-TTGTTATGAAATTCTATA----TGCGTATT-GATGATGA
SEQ2308     ATATCATTCCCAAAGGTGA-TTGTTATGAA-CTCTATA----TGCGTATT-GATGATGA
SEQ2309     ATATCATTCCCAAAGGTGA-TTGTTATGAA-CTCTATA----TGCGTATT-GATGATGA
SEQ2310     ----------------TGACTTGTTATGAAACTCTATA----TGCGTATTTGATGATGA
SEQ2311     ATATCATTCCCAAAGGTGA-TTGTTATGAA-TTCTATA----TGCGTATT-GATGATGA

SEQ2301     AA-GGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA
SEQ2302     AA-GGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA
SEQ2303     AA-GGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA
SEQ2304     AA-GGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA
SEQ2305     AA-GGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA
SEQ2306     AA-GGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA
SEQ2307     AA-GGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA
SEQ2308     AA-GGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA
SEQ2309     AA-GGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA
SEQ2310     AAAGGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA
SEQ2311     AA-GGCGGTTTATTGATGTTTTTGAGTTTAATAGGATGGCTAGTCTTATTAGTCACTTTA

SEQ2301     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
SEQ2302     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
SEQ2303     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
SEQ2304     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
SEQ2305     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
SEQ2306     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
SEQ2307     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
SEQ2308     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
SEQ2309     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
SEQ2310     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
SEQ2311     AATTTGTGGCAGGCATGAACGTTGGAGAAAAAAGACGAAGTCAATTAGGTTCTTGTGACT
```

Table 23: Comparative Sequences relating to SAG0163 (competence protein Cg1A)

SEQ2301	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTG
SEQ2302	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTG
SEQ2303	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTG
SEQ2304	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTG
SEQ2305	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTG
SEQ2306	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTG
SEQ2307	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCAAGTGTGGGAGATTATCGTG
SEQ2308	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTG
SEQ2309	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCGAGTGTGGGAGATTATCGTG
SEQ2310	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCAAGTGTGGGAGATTATCGTG
SEQ2311	ATGAACTGTCAGAGGGAAGACTGGTTTCATTACGACTATCAAGTGTGGGAGATTATCGTG
SEQ2301	GTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2302	GTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2303	GTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2304	GTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2305	GTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2306	GTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2307	GTCAAGAATCTTTAGTTATTCGTACTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2308	GTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2309	GTCAAGAATCTTTAGTTATTCGTATTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2310	GTCAAGAATCTTTAGTTATTCGTACTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2311	GTCAAGAATCTTTAGTTATTCGTACTTTGTATTCAGGTCATCAGGACTTAAAATATTGGT
SEQ2301	TTGATAATATAAAGCAAATGAAGGAAGTACTGGGTACAAGAGGGCTATATCTTTTTTCCG
SEQ2302	TTGATAATATAAAGCAAATGAAGGAAGTACTGGGTACAAGAGGGCTATATCTTTTTTCCG
SEQ2303	TTGATAATATAAAGCAAATGAAGGAAGTACTGGGTATAAGAGGGCTATATCTTTTTTCCG
SEQ2304	TTGATAATATAAAGCAAATGAAGGAAGTACTGGGTATAAGAGGGCTATATCTTTTTTCCG
SEQ2305	TTGATAATATAAAGCAAATGAAGGAAGTACTGGGTATAAGAGGGCTATATCTTTTTTCCG
SEQ2306	TTGATAATATAAAGCAAATGAAGGAAGTACTGGGTACAAGAGGGCTATATCTTTTTTCCG
SEQ2307	TTGATAATATAAAGTAAATGAAGGAAGTACTGTGTGCAAGAGGGCTATATCTTTTTTCCG
SEQ2308	TTGATAATATAAAGCAAATGAAGGAAGTACTGGGTATAAGAGGGCTATATCTTTTTTCCG
SEQ2309	TTGATAATATAAAGCAAATGAAGGAAGTACTGGGTATAAGAGGGCTATATCTTTTTTCCG
SEQ2310	TTGATAATATAAAGTAAATGAAGGAAGTACTGTGTGCAAGAGGGCTATATCTTTTTTCCG
SEQ2311	TTGATAATATAAAGCAAATGAAGGAAGTACTGTGTGCAAGAGGGCTATATCTTTTTTCCG
SEQ2301	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA
SEQ2302	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA
SEQ2303	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA
SEQ2304	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA
SEQ2305	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA
SEQ2306	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA
SEQ2307	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA
SEQ2308	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA
SEQ2309	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA
SEQ2310	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA
SEQ2311	GCCCTGTGGGGAGTGGTAAAACAACTCTCATGTATCAATTAGCTTCAGAAGTATTTAAAA

Table 23: Comparative Sequences relating to SAG0163 (competence protein Cg1A)

```
SEQ2301          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC
SEQ2302          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC
SEQ2303          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC
SEQ2304          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC
SEQ2305          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC
SEQ2306          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC
SEQ2307          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC
SEQ2308          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC
SEQ2309          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC
SEQ2310          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC
SEQ2311          ATAAGCAAATTATCACGATTGAAGATCCGGTAGAAATCAAGAATGACAAGATGTTACAAC

SEQ2301          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
SEQ2302          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
SEQ2303          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
SEQ2304          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
SEQ2305          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
SEQ2306          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
SEQ2307          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
SEQ2308          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
SEQ2309          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
SEQ2310          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC
SEQ2311          TCCAATTGAATGAGGATATTGGAATGACTTATGATGCTTTAATCAAACTGTCTTTACGGC

SEQ2301          ATCGTCCAGATATTTTAATTATCGGAGAGAT-TAGAGATCAAGCGACGGCCCGTGCTGTT
SEQ2302          ATCGTCCAGATATTTTAATTATCGGAGAGAT-TAGAGATCAAGCGACGGCTCGTGCTGTT
SEQ2303          ATCGTCCAGATATTTTAATTATCGGAGAGAT-TAGAGATCAAGCGACGGCCCGTGCTGTT
SEQ2304          ATCGTCCAGATATTTTAATTATCGGAGAGAT-TAGAGATCAAGCGACGGCCCGTGCTGTT
SEQ2305          ATCGTCCAGATATTTTAATTATCGGAGAGAT-TAGAGATCAAGCGACGGCCCGTGCTGTT
SEQ2306          ATCGTCCAGATATTTTAATTATCGGAGAGAT-TAGAGATCAAGCGACGGCCCGTGCTGTT
SEQ2307          ATCGTCCAGATATTTTAATTATCGGAGAGAT-TAGAGATCAAGCGACGGCCCGTGCTGTT
SEQ2308          ATCGTCCAGATATTTTAATTATCGGAGAGAAATAGAGATCAAGCGACGGCCCGTGCTGTT
SEQ2309          ATCGTCCAGATATTTTAATTATCGGAGAGAT-TAGAGATCAAGCGACGGCCCGTGCTGTT
SEQ2310          ATCGTCCAGATATTTTAATTATCGGAGAGAT-TAGAGATCAAGCGACGGCCCGTGCTGTT
SEQ2311          ATCGTCCAGATATTTTAATTATCGGAGAGAT-TAGAGATCAAGCGACGGCCCGTGCTGTT

SEQ2301          ATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCATGCTAAAAGTATTTCC
SEQ2302          ATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCATGCTAAAAGTATTCCC
SEQ2303          ATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCATGCTAAAAGTATTCCC
SEQ2304          ATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCATGCTAAAAGTATTCCC
SEQ2305          ATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCATGCTAAAAGTATTCCC
SEQ2306          ATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCATGCTAAAAGTATTTCC
SEQ2307          ATTCGTGCAAGTTTAACGGGAGTAATGGTTTTTTCTACTATTCATGCTAAAAGTATTCCC
SEQ2308          ATTCGTGCAAGTTTAACGGGAGTGATGTTTTTTTCTACTATTCATGCTAAAAGTATTCCC
SEQ2309          ATTCGTGCAAGTTTAACGGGAGTGATGGTTTTTTCTACTATTCATGCTAAAAGTATTCCC
SEQ2310          ATTCGTGCAAGTTTAACGGGAGTAATGGTTTTTTCTACTATTCATGCTAAAAGTATTCCC
SEQ2311          ATTCGTGCAAGTTTAACGGGAGTAATGGTTTTTTCTACTATTCATGCTAAAAGTATTCCC

SEQ2301          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA
SEQ2302          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA
SEQ2303          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA
SEQ2304          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA
SEQ2305          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA
SEQ2306          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA
SEQ2307          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA
SEQ2308          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA
SEQ2309          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA
SEQ2310          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA
SEQ2311          GGAGTCTATGATAGGCTTATAGAATTAGGGGTTAACTATCAAGAGTTAGAAAATAGTCTA

SEQ2301          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGT
SEQ2302          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAAGT
SEQ2303          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGT
SEQ2304          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGT
SEQ2305          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGT
SEQ2306          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGT
SEQ2307          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAAGT
```

Table 23: Comparative Sequences relating to SAG0163 (competence protein Cg1A)

```
SEQ2308          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGT
SEQ2309          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAGGT
SEQ2310          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAAGT
SEQ2311          AAATTAATAGCATATCAACGTTTAATTGGAGGAGGAAGCCTAATTGACTTTGAGACAAGT

SEQ2301          AACTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA
SEQ2302          AACTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA
SEQ2303          AATTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA
SEQ2304          AATTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA
SEQ2305          AATTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA
SEQ2306          AACTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA
SEQ2307          AACTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA
SEQ2308          AATTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA
SEQ2309          AATTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA
SEQ2310          AACTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA
SEQ2311          AACTTTAAAAAACACTCATCAGACAAGTGGAATAGACAAGTGGATATCTTGGCTGAAGAA

SEQ2301          GGACATATCAGTAAGAAACAGGCACAAGT-CGAAAAAATTATCCCTCAAGAAACAACGGA
SEQ2302          GGATATATCAGTAAGAAACAGGCACAAGT-CGAAAAAATTATCCCTCAAGAAACAACGGA
SEQ2303          GGACATATCAGTAAGAAACAGGCACAAGT-CGAAAAAATTATCCCTCAAGAAACAACGGA
SEQ2304          GGACATATCAGTAAGAAACAGGCACAAGTGCGAAAAAATTATCCCTCAAGAAACAACGGA
SEQ2305          GGACATATCAGTAAGAAACAGGCACAAGT-CGAAAAAATTATCCCTCAAGAAACAACGGA
SEQ2306          GGACATATCAGTAAGAAACAGGCACAAGT-CGAAAAAATTATCCCTCAAGAAACAACGGA
SEQ2307          GGACATATCAGTAAGAAACAGGCACAAGT-CGAAAAAATTATCCCTCAAGAAACAACGGA
SEQ2308          GGACATATCAGTAAGAAACAGGCACAAGT-CGAAAAAATTATCCCTCAAGAAACAACGGA
SEQ2309          GGACATATCAGTAAGAAACAGGCACAAGT-CGAAAAAATTATCCCTCAAGAAACAACGGA
SEQ2310          GGACATATCAGTAAGAAACAGGCACAAGT-CGAAAAAATTATCCCTCAAGAAACAACGGA
SEQ2311          GGACATATCAGTAAGAAACAGGCACAAGT-CGAAAAAATTATCCCTCAAGAAACAACGGA

SEQ2301          AAGTAGTCCAACTTTT-----------------------------------
SEQ2302          AAGTAGTCCAACTTTT-----------------------------------
SEQ2303          AAGTAGTCCAACTTTT-----------------------------------
SEQ2304          AAGTAGTCCAACTTTT-----------------------------------
SEQ2305          AAGTAGTCCAACTTTT-----------------------------------
SEQ2306          AAGTAGTCCAACTTTT-----------------------------------
SEQ2307          AAGTAGTCCAACTTTT-----------------------------------
SEQ2308          AAGTAGTCCAACTTTT-----------------------------------
SEQ2309          AAGTAGTCCAACTTTT-----------------------------------
SEQ2310          AAGTAGTCCAACTTTT-----------------------------------
SEQ2311          AAGTAGTCCAACTTTT
```

>SEQ ID NO 2350:63_090 frame: 2
AVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFVAGMNVGEKRRS
QLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDNIKQMKEVLGTR
GLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQLNEDIGMTYDAL
IKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSISGVYDRLIELGVNYQ
ELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHISKKQAQVEKII
PQETTESSPTF

>SEQ ID NO 2351:63_1169NT frame: 3
.LL.NLYYCVFDDERRFIDVFEFNRMASLISHFKFVAGMNVGEKRRSQLGSCDYELSEGR
LVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDNIKQMKEVLGTRGLYLFSGPVGSGK
TTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQLNEDIGMTYDALIKLSLRHRPDILI
IGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVYDRLIELGVNYQELENSLKLIAYQR
LIGGGSLIDFETSNFKKHSSDKWNRQVDILAEEGYISKKQAQVEKIIPQETTESSPTF

>SEQ ID NO 2352:63_18RS21 frame: 1
VQSLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
IKQMKEVLGIRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
NEDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVY
DRLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SKKQAQVEKIIPQETTESSPTF

>SEQ ID NO 2353: 63_2603 frame: 1
DIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFVAGMNVGEKRRSQLGSCDY
ELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDNIKQMKEVLGIRGLYLFSG

Table 23: Comparative Sequences relating to SAG0163 (competence protein Cg1A)

PVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQLNEDIGMTYDALIKLSLRH
RPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVYDRLIELGVNYQELENSLK
LIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHISKKQAQVRKNYPSRNNGK
.SNF
>SEQ ID NO 2354:63_A909 frame: 1
VQSLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
IKQMKEVLGIRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
NEDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVY
DRLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SKKQAQVEKIIPQETTESSPTF

>SEQ ID NO 2355:63_CJB110 frame: 1
VQSLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
IKQMKEVLGTRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
NEDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSISGVY
DRLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SKKQAQVEKIIPQETTESSPTF

>SEQ ID NO 2356:63_CJB110 frame: 1
VQSLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
IKQMKEVLGTRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
NEDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSISGVY
DRLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SKKQAQVEKIIPQETTESSPTF
>SEQ ID NO 2357: 63_H36B frame: 1
SLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFVAG
MNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDNIK
QMKEVLGIRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQLNE
DIGMTYDALIKLSLRHRPDILIIGEK

>SEQ ID NO 2358:63_JM9130013 frame: 1
VQSLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
IKQMKEVLGIRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
NEDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVY
DRLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SKKQAQVEKIIPQETTESSPTF

>SEQ ID NO 2359:63_M732 frame: 3
TCYETLYAYLMMKRRFIDVFEFNRMASLISHFKFVAGMNVGEKRRSQLGSCDYELSEGRL
VSLRLSSVGDYRGQESLVIRTLYSGHQDLKYWFDNIK.MKEVLCARGLYLFSGPVGSGKT
TLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQLNEDIGMTYDALIKLSLRHRPDILII
GEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVYDRLIELGVNYQELENSLKLIAYQRL
IGGGSLIDFETSNFKKHSSDKWNRQVDILAEEGHISKKQAQVEKIIPQETTESSPTF

>SEQ ID NO 2360:63_M781 frame: 3
VEVNAQDIYIIPKGDCYEFYMRIDDERRFIDVFEFNRMASLISHFKFVAGMNVGEKRRSQ
LGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRTLYSGHQDLKYWFDNIKQMKEVLCARG
LYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQLNEDIGMTYDALI
KLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVYDRLIELGVNYQE
LENSLKLIAYQRLIGGGSLIDFETSNFKKHSSDKWNRQVDILAEEGHISKKQAQVEKIIP
QETTESSPTF

>SEQ ID NO 2361:63_COH1 frame: 3
VIVMKFYMRIDDERRFIDVFEFNRMASLISHFKFVAGMNVGEKRRSQLGSCDYELSEGRL
VSLRLSSVGDYRGQESLVIRTLYSGHQDLKYWFDNIK

Table 23: Comparative Sequences relating to SAG0163 (competence protein Cg1A)

```
SEQ2350          ----------AVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
SEQ2351          -------------------------LLNLYYCVFDDERRFIDVFEFNRMASLISHFKFV
SEQ2352          QSLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
SEQ2353          -----------------DIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
SEQ2354          QSLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
SEQ2355          QSLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
SEQ2356          QSLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
SEQ2357          -SLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
SEQ2358          QSLAKQVIHQAVEVNAQDIYIIPKGDCYELYMRIDDERRFIDVFEFNRMASLISHFKFV
SEQ2359          ------------------------TCYETLYAYLMMKRRFIDVFEFNRMASLISHFKFV
SEQ2360          -----------VEVNAQDIYIIPKGDCYEFYMRIDDERRFIDVFEFNRMASLISHFKFV
SEQ2361          ------------------------VIVMKFYMRIDDERRFIDVFEFNRMASLISHFKFV

SEQ2350          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
SEQ2351          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
SEQ2352          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
SEQ2353          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
SEQ2354          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
SEQ2355          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
SEQ2356          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
SEQ2357          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
SEQ2358          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRILYSGHQDLKYWFDN
SEQ2359          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRTLYSGHQDLKYWFDN
SEQ2360          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRTLYSGHQDLKYWFDN
SEQ2361          AGMNVGEKRRSQLGSCDYELSEGRLVSLRLSSVGDYRGQESLVIRTLYSGHQDLKYWFDN

SEQ2350          IKQMKEVLGTRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2351          IKQMKEVLGTRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2352          IKQMKEVLGIRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2353          IKQMKEVLGIRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2354          IKQMKEVLGIRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2355          IKQMKEVLGTRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2356          IKQMKEVLGTRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2357          IKQMKEVLGIRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2358          IKQMKEVLGIRGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2359          IK-MKEVLCARGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2360          IKQMKEVLCARGLYLFSGPVGSGKTTLMYQLASEVFKNKQIITIEDPVEIKNDKMLQLQL
SEQ2361          IK----------------------------------------------------------

SEQ2350          EDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSISGVY
SEQ2351          EDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVY
SEQ2352          EDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVY
SEQ2353          EDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVY
SEQ2354          EDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVY
SEQ2355          EDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSISGVY
SEQ2356          EDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSISGVY
SEQ2357          EDIGMTYDALIKLSLRHRPDILIIGEK--------------------------------
SEQ2358          EDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVY
SEQ2359          EDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVY
SEQ2360          EDIGMTYDALIKLSLRHRPDILIIGEIRDQATARAVIRASLTGVMVFSTIHAKSIPGVY
SEQ2361          -----------------------------------------------------------

SEQ2350          RLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SEQ2351          RLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETSNFKKHSSDKWNRQVDILAEEGYI
SEQ2352          RLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SEQ2353          RLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SEQ2354          RLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SEQ2355          RLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SEQ2356          RLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SEQ2357          -----------------------------------------------------------
SEQ2358          RLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETGNFKKHSSDKWNRQVDILAEEGHI
SEQ2359          RLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETSNFKKHSSDKWNRQVDILAEEGHI
SEQ2360          RLIELGVNYQELENSLKLIAYQRLIGGGSLIDFETSNFKKHSSDKWNRQVDILAEEGHI
SEQ2361          -----------------------------------------------------------
```

Table 23: Comparative Sequences relating to SAG0163 (competence protein Cg1A)

```
SEQ2350          KKQAQVEKIIPQETTESSPTF
SEQ2351          KKQAQVEKIIPQETTESSPTF
SEQ2352          KKQAQVEKIIPQETTESSPTF
SEQ2353          KKQAQVRKNYPSRNNGKSNF-
SEQ2354          KKQAQVEKIIPQETTESSPTF
SEQ2355          KKQAQVEKIIPQETTESSPTF
SEQ2356          KKQAQVEKIIPQETTESSPTF
SEQ2357          ---------------------
SEQ2358          KKQAQVEKIIPQETTESSPTF
SEQ2359          KKQAQVEKIIPQETTESSPTF
SEQ2360          KKQAQVEKIIPQETTESSPTF
SEQ2361          ---------------------
```

Figure 24: Comparative Sequences relating to SAG0290 (ABC transporter, substrate-binding protein)

SEQ ID NO. 2401: SAG0290 FROM THE 1169NT1 GBS TYPE V STRAIN (REVERSE COMPLEMENT)

GTATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCACCATTTACTTA
TCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACA
AAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCA
GCTAATGATTTTTCATACAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGACCCTATATCCCGTTCAAA
TTATGCCGTAGTAGGGAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACAGAAG
TTTTATCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAACCAATA
AAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGGAAAATTGA
CTTTATCCTATATGATGCCATTTCATCTGACTATATTGTAAAAGATCAATCATTAAACTTAAGCGTTTCTC
CTTTGAAAGGTAAAATTGGTAATAATAAGGATGGATTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGT
AAAACTCTACAGAAATTTATAAATAAGCGTATTAAAGTTTTGAAAGAAGATGGTACTTTGGCACGTTTAAG
TAAACAATATTTCGGTGGAGATTACGTTTCAAACATTGATAAA

SEQ ID NO. 2402: SAG0290 FROM THE 18RS21 GBS TYPE II STRAIN (REVERSE COMPLEMENT)

GTATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCACCATTTACTTA
TRAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACA
AAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCA
GCTAATGATTTTTCATACAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAA
TTATGCCGTAGTAGGGAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAG
TTTTATCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAACCAATA
AAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGGAAAATTGA
CTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAAGACCAATCATTAAACTTAAGCGTTTCTC
CTTTGAAAGGTAAAATTGGTAATAATAAGGATGGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGT
AAAACTCTACAGAAATTTATAAATAAGCGTATTAAAGTTTTGAAAGAAAATGGTACTTTGGCACGTTTAAG
TAAACAATATTTCGGTGGAGATTACGTTTCAAACATTGATAAA

SEQ ID NO. 2403: SAG0290 FROM THE 2603 V/R GBS TYPE V STRAIN (REVERSE COMPLEMENT)

ATTCAAAGGTTATGATGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGA
CAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCA
TACAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGG
GAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAGTTTTATCTGGCGTTA
ACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAACCAATAAAAATCAAATATGTT
TCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGGAAAATTGACTTTATCCTATATGA
TGCCATTTCATCCGACTATATTGTAAAAGACCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAA
TTGGTAATAATAAGGATGGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAG

SEQ ID NO. 2404: SAG0290 FROM THE 090 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)

GTATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCACCATTTACTTA
TCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACA
AAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCA
GCTAATGATTTTTCATACAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAA
TTATGCCGTAGTAGGGAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAG
TTTTATCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAACCAATA
AAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGGAAAATTGA
CTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAAGACCAATCATTAAACTTAAGCGTTTCTCCT
TTGAAAGGTAAAATTGGTAATAATAAGGATGGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTA
CAGAAATTTATAAATAAGCGTATTAAAGTTTTGAAAGAAAATGGTACTTTGGCACGTTTAAGTAAACAATATTTCGGT
GGAGATTACGTTTCAAACATTGATAAA

Figure 24: Comparative Sequences relating to SAG0290
(ABC transporter, substrate-binding protein)

SEQ ID NO. 2405: SAG0290 FROM THE A909 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)
GTATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCACCATTTACTTA
TCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACA
AAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCA
GCTAATGATTTTTCATACAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAA
TTATGCCGTAGTAGGGAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAG
TTTTATCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATNNTAATAAAAAACCANTA
AAAATNAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGAAAATTGA
CTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAAGACCAATCATTAAACTTAAGCGTTTCTC
CTTTGAAAGGTAAAATTGGTAATAATAAGGATGGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGT
AAAACTCTACAGAAATTTATAAATAAGCGT

SEQ ID NO. 2406: SAG0290 FROM THE CJB110 GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)
GTATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCACCATTTACTTA
TCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACA
AAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCA
GCTAATGATTTTTCATACAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAA
TTATGCCGTAGTAGGGAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAG
TTTTATCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAACCAATA
AAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGAAAATTGA
CTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAAGACCAATCATTAAACTTAAGCGTTTCTC
CTTTGAAAGGTAAAATTGGTAATAATAAGGATGGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGT
AAAACTCTACAGAAATTTATAAATAAGCGTATTAAAGTTTTGAAAGAAAATGGTACTTTGGCACGTTTAAG
TAAACAATATTTCGGTGGAGATTACGTTTCAAACATTGATAAA

SEQ ID NO. 2407: SAG0290 FROM THE COH1 GBS TYPE III STRAIN (REVERSE COMPLEMENT)
GTATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCACCATTTACTTA
TCAAAAAGACGGGAAATTCAAAGGTTATGACGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACA
AAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCA
GCTAATGATTTTTCATATAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAA
TTATGCCGTAGTAGGGAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACAGAAG
TTTTATCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAACCAATA
AAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGAAAAATTGA
CTTTATCCTATATGATGCCATTTCATCTGACTATATTGTAAAAGATCAATCATTAAACTTAAGCGTTTCTC
CTTTGAAAGGTAAAATTGGTAATAATAAGGATGGATTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGT
AAAACTCTACAGAAATTTATAAATAAGCGTATTAAAGTTTTGAAAGAAGATGGTACTTTGGCACGTTTAAG
TAAACAATATTTCGGTGGAGATTACGTTTCAAACATTGATAAA

SEQ ID NO. 2408: SAG0290 FROM THE H36b GBS TYPE Ib STRAIN (REVERSE COMPLEMENT)
GTATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCACCATTTACTTA
TCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACA
AAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCA
GCTAATGATTTTTCATACAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAA
TTATGCCGTAGTAGGGAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAG
TTTTATCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAACCAATA
AAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGAAAATTGA
CTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAAGACCAATCATTAAACTTAAGCGTTTCTC
CTTTGAAAGGTAAAATTGGTAATAATAAGGATGGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGT
AAAACTCTACAGAAATTTATAAATAAGCGTATTAAAGTTTTGAAAGAAAATGGTACTTTGGCACGTTTAAG
TAAACAATATTTCGGTGGAGATTACGTTTCAAACATTGATAAA

Figure 24: Comparative Sequences relating to SAG0290 (ABC transporter, substrate-binding protein)

SEQ ID NO. 2409: SAG0290 FROM THE JM9130013 GBS STRAIN VIII (REVERSE COMPLEMENT)

GTATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCACCATTTACTTA
TCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACA
AAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCA
GCTAATGATTTTTCATACAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAA
TTATGCCGTAGTAGGGAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAG
TTTTATCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAACCAATA
AAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGAAAATTGA
CTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAAGACCAATCATTAAACTTAAGCGTTTCTC
CTTTGAAAGGTAAAATTGGTAATAATAAGGATGGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGT
AAAACTCTACAGAAATTTATAAATAAGCGTAATAAAGTTTTGAAAGAAAATGGTA

SEQ ID NO. 2410: SAG0290 FROM THE M732 GBS TYPE III STRAIN (REVERSE COMPLEMENT)

GTATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCACCATTTACTTA
TCAAAAAGACGGGAAATTCAAAGGTTATGACGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACA
AAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCA
GCTAATGATTTTTCATATAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAA
TTATGCCGTAGTAGGGAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACAGAAG
TTTTATCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAACCAATA
AAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGAAAAATTGA
CTTTATCCTATATGATGCCATTTCATCTGACTATATTGTAAAAGATCAATCATTAAACTTAAGCGTTTCTC
CTTTGAAAGGTAAAATTGGTAATAATAAGGATGGATTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGT
AAAACTCTACAGAAATTTATAAATAAGCGTATTAAAGTTTTGAAAGAAGATGGTACTTTGGCACGTTTAAG
TAAACAATATTTCGGTGGAGATTACGTTTCAAACATTGATAAA

SEQ ID NO. 2411: SAG0290 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)

GTATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCACCATTTACTTA
TCAAAAAGACGGGAAATTCAAAGGTTATGACGTTGATGTTGTCAAAGCTGTTTTTAAAGGTAGTAAGTACA
AAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCAACAGGTATTGATGCAGGGAAATTTGATTTATCA
GCTAATGATTTTTCATATAATAAAGAAAGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAA
TTATGCCGTAGTAGGGAAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACAGAAG
TTTTATCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAACCAATA
AAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATTGAGAGTGGAAAAATTGA
CTTTATCCTATATGATGCCATTTCATCTGACTATATTGTAAAAGATCAATCATTAAACTTAAGCGTTTCTC
CTTTGAAAGGTAAAATTGGTAATAATAAGGATGGATTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGT
AAAACTCTACAGAAATTTATAAATAAGCGTATTAAAGTTTTGAAAGAAGATGGTACTTTGGCACGTTTAAG
TAAACAATATTTCGGTGGAGATTACGTTTCAAACATTGATAAA

```
SEQ2401         TATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCA
SEQ2402         TATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCA
SEQ2403         -----------------------------------------------------------
SEQ2404         TATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCA
SEQ2405         TATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCA
SEQ2406         TATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCA
SEQ2407         TATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCA
SEQ2408         TATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCA
SEQ2409         TATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCA
SEQ2410         TATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCA
SEQ2411         TATCAGTTCAGGCGTCAGAGAAAGTAGAACTTAAAGTAGCTACAGATTCTGACACGGCA


SEQ2401         CATTTACTTATCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCT
SEQ2402         CATTTACTTATRAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCT
SEQ2403         -------------------------ATTCAAAGGTTATGATGTTGATGTTGTCAAAGCT
SEQ2404         CATTTACTTATCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCT
SEQ2405         CATTTACTTATCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCT
SEQ2406         CATTTACTTATCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCT
SEQ2407         CATTTACTTATCAAAAAGACGGGAAATTCAAAGGTTATGACGTTGATGTTGTCAAAGCT
SEQ2408         CATTTACTTATCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCT
SEQ2409         CATTTACTTATCAAAAAGACGGGAAATTCAAAGGTTATGATGTTGATGTTGTCAAAGCT
```

Figure 24: Comparative Sequences relating to SAG0290 (ABC transporter, substrate-binding protein)

SEQ2410 CATTTACTTATCAAAAAGACGGGAAATTCAAAGGTTATGACGTTGATGTTGTCAAAGCT
SEQ2411 CATTTACTTATCAAAAAGACGGGAAATTCAAAGGTTATGACGTTGATGTTGTCAAAGCT

SEQ2401 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA
SEQ2402 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA
SEQ2403 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA
SEQ2404 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA
SEQ2405 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA
SEQ2406 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA
SEQ2407 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA
SEQ2408 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA
SEQ2409 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA
SEQ2410 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA
SEQ2411 GTTTTTAAAGGTAGTAAGTACAAAGTAACCTTCAAGACAGTTCCTTTTGATACTATTTCA

SEQ2401 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATACAATAAAGAA
SEQ2402 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATACAATAAAGAA
SEQ2403 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATACAATAAAGAA
SEQ2404 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATACAATAAAGAA
SEQ2405 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATACAATAAAGAA
SEQ2406 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATACAATAAAGAA
SEQ2407 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATATAATAAAGAA
SEQ2408 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATACAATAAAGAA
SEQ2409 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATACAATAAAGAA
SEQ2410 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATATAATAAAGAA
SEQ2411 ACAGGTATTGATGCAGGGAAATTTGATTTATCAGCTAATGATTTTTCATATAATAAAGAA

SEQ2401 AGAGCAGAAAAATATCTCTTCTCAGACCCTATATCCCGTTCAAATTATGCCGTAGTAGGG
SEQ2402 AGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGGG
SEQ2403 AGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGGG
SEQ2404 AGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGGG
SEQ2405 AGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGGG
SEQ2406 AGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGGG
SEQ2407 AGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGGG
SEQ2408 AGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGGG
SEQ2409 AGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGGG
SEQ2410 AGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGGG
SEQ2411 AGAGCAGAAAAATATCTCTTCTCAGATCCTATATCCCGTTCAAATTATGCCGTAGTAGGG

SEQ2401 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACAGAAGTTTTA
SEQ2402 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAGTTTTA
SEQ2403 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAGTTTTA
SEQ2404 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAGTTTTA
SEQ2405 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAGTTTTA
SEQ2406 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAGTTTTA
SEQ2407 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACAGAAGTTTTA
SEQ2408 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAGTTTTA
SEQ2409 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACCGAAGTTTTA
SEQ2410 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACAGAAGTTTTA
SEQ2411 AAGAAGGGGAGCCATTACAAATCATTAAGTGACCTCTCTGGAAAATCAACAGAAGTTTTA

SEQ2401 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAA
SEQ2402 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAA
SEQ2403 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAA
SEQ2404 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAA
SEQ2405 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATNNTAATAAAAAA
SEQ2406 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAA
SEQ2407 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAA
SEQ2408 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAA
SEQ2409 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAA
SEQ2410 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAA
SEQ2411 TCTGGCGTTAACTATGCACAGGTTCTAGAAAATTGGAATAAAAATCATCCTAATAAAAAA

SEQ2401 CCAATAAAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT
SEQ2402 CCAATAAAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT
SEQ2403 CCAATAAAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT
SEQ2404 CCAATAAAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT
SEQ2405 CCANTAAAAATNAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT
SEQ2406 CCAATAAAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT
SEQ2407 CCAATAAAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT
SEQ2408 CCAATAAAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT
SEQ2409 CCAATAAAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT

Figure 24: Comparative Sequences relating to SAG0290 (ABC transporter, substrate-binding protein)

```
SEQ2410        CCAATAAAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT
SEQ2411        CCAATAAAAATCAAATATGTTTCTGGGACAACTGGTGTTACTAGCAGATTAAAAAATATT

SEQ2401        GAGAGTGGGAAAATTGACTTTATCCTATATGATGCCATTTCATCTGACTATATTGTAAAA
SEQ2402        GAGAGTGGGAAAATTGACTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAA
SEQ2403        GAGAGTGGGAAAATTGACTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAA
SEQ2404        GAGAGTGGGAAAATTGACTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAA
SEQ2405        GAGAGTGGGAAAATTGACTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAA
SEQ2406        GAGAGTGGGAAAATTGACTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAA
SEQ2407        GAGAGTGGAAAAATTGACTTTATCCTATATGATGCCATTTCATCTGACTATATTGTAAAA
SEQ2408        GAGAGTGGGAAAATTGACTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAA
SEQ2409        GAGAGTGGGAAAATTGACTTTATCCTATATGATGCCATTTCATCCGACTATATTGTAAAA
SEQ2410        GAGAGTGGAAAAATTGACTTTATCCTATATGATGCCATTTCATCTGACTATATTGTAAAA
SEQ2411        GAGAGTGGAAAAATTGACTTTATCCTATATGATGCCATTTCATCTGACTATATTGTAAAA

SEQ2401        GATCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT
SEQ2402        GACCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT
SEQ2403        GACCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT
SEQ2404        GACCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT
SEQ2405        GACCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT
SEQ2406        GACCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT
SEQ2407        GATCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT
SEQ2408        GACCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT
SEQ2409        GACCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT
SEQ2410        GATCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT
SEQ2411        GATCAATCATTAAACTTAAGCGTTTCTCCTTTGAAAGGTAAAATTGGTAATAATAAGGAT

SEQ2401        GGATTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTACAGAAATTTATA
SEQ2402        GGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTACAGAAATTTATA
SEQ2403        GGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAG-----------------------
SEQ2404        GGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTACAGAAATTTATA
SEQ2405        GGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTACAGAAATTTATA
SEQ2406        GGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTACAGAAATTTATA
SEQ2407        GGATTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTACAGAAATTTATA
SEQ2408        GGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTACAGAAATTTATA
SEQ2409        GGACTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTACAGAAATTTATA
SEQ2410        GGATTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTACAGAAATTTATA
SEQ2411        GGATTAGAATACCTCCTTTTACCAAAAGATAAAAAAGGTAAAACTCTACAGAAATTTATA

SEQ2401        ATAAGCGTATTAAAGTTTTGAAAGAAGATGGTACTTTGGCACGTTTAAGTAAACAATAT
SEQ2402        ATAAGCGTATTAAAGTTTTGAAAGAAAATGGTACTTTGGCACGTTTAAGTAAACAATAT
SEQ2403        -----------------------------------------------------------
SEQ2404        ATAAGCGTATTAAAGTTTTGAAAGAAAATGGTACTTTGGCACGTTTAAGTAAACAATAT
SEQ2405        ATAAGCGT---------------------------------------------------
SEQ2406        ATAAGCGTATTAAAGTTTTGAAAGAAAATGGTACTTTGGCACGTTTAAGTAAACAATAT
SEQ2407        ATAAGCGTATTAAAGTTTTGAAAGAAGATGGTACTTTGGCACGTTTAAGTAAACAATAT
SEQ2408        ATAAGCGTATTAAAGTTTTGAAAGAAAATGGTACTTTGGCACGTTTAAGTAAACAATAT
SEQ2409        ATAAGCGTAATAAAGTTTTGAAAGAAAATGGTA--------------------------
SEQ2410        ATAAGCGTATTAAAGTTTTGAAAGAAGATGGTACTTTGGCACGTTTAAGTAAACAATAT
SEQ2411        ATAAGCGTATTAAAGTTTTGAAAGAAGATGGTACTTTGGCACGTTTAAGTAAACAATAT

SEQ2401        TCGGTGGAGATTACGTTTCAAACATTGATAAA---------------------------
SEQ2402        TCGGTGGAGATTACGTTTCAAACATTGATAAA---------------------------
SEQ2403        -----------------------------------------------------------
SEQ2404        TCGGTGGAGATTACGTTTCAAACATTGATAAA---------------------------
SEQ2405        -----------------------------------------------------------
SEQ2406        TCGGTGGAGATTACGTTTCAAACATTGATAAA---------------------------
SEQ2407        TCGGTGGAGATTACGTTTCAAACATTGATAAA---------------------------
SEQ2408        TCGGTGGAGATTACGTTTCAAACATTGATAAA---------------------------
SEQ2409        -----------------------------------------------------------
SEQ2410        TCGGTGGAGATTACGTTTCAAACATTGATAAA---------------------------
SEQ2411        TCGGTGGAGATTACGTTTCAAACATTGATAAAGTRCMARATVSTNCSRATNGTSAGABC
```

Figure 24: Comparative Sequences relating to SAG0290 (ABC transporter, substrate-binding protein)

```
SEQ2401         -----------------------
SEQ2402         -----------------------
SEQ2403         -----------------------
SEQ2404         -----------------------
SEQ2405         -----------------------
SEQ2406         -----------------------
SEQ2407         -----------------------
SEQ2408         -----------------------
SEQ2409         -----------------------
SEQ2410         -----------------------
SEQ2411         RANSRTRSTBSTRATBNDNGRTN
```

```
>SEQ ID NO 2450: 8_1169NT frame: 1
VSVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKEDGTLARLSKQY
FGGDYVSNIDK

>SEQ ID NO 2451:8_18RS21 frame: 1
VSVQASEKVELKVATDSDTAPFTYXKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKENGTLARLSKQY
FGGDYVSNIDK

>SEQ ID NO 2452:8_2603 frame: 2
FKGYDVDVVKAVFKGSKYKVTFKTVPFDTISTGIDAGKFDLSANDFSYNKERAEKYLFSD
PISRSNYAVVGKKGSHYKSLSDLSGKSTEVLSGVNYAQVLENWNKNHPNKKPIKIKYVSG
TTGVTSRLKNIESGKIDFILYDAISSDYIVKDQSLNLSVSPLKGKIGNNKDGLEYLLLPK
DKK

>SEQ ID NO 2453:8_090 frame: 1
VSVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKENGTLARLSKQY
FGGDYVSNIDK

>SEQ ID NO 2454:8_A909 frame: 1
VSVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SGVNYAQVLENWNKNHXNKKPXKXKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKR

>SEQ ID NO 2455: 8_CJB110 frame: 1
VSVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKENGTLARLSKQY
FGGDYVSNIDK

>SEQ ID NO 2456: 8_COH1 frame: 1
VSVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKEDGTLARLSKQY
FGGDYVSNIDK

>SEQ ID NO 2457:8_H36B frame: 1
VSVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKENGTLARLSKQY
FGGDYVSNIDK
```

Figure 24: Comparative Sequences relating to SAG0290
(ABC transporter, substrate-binding protein)

>SEQ ID NO 2458:8_JM9130013 frame: 1
VSVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRNKVLKENG

>SEQ ID NO 2459:8_M732 frame: 1
VSVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKEDGTLARLSKQY
FGGDYVSNIDK

>SEQ ID NO 2460:8_M781 frame: 1
VSVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKEDGTLARLSKQY
FGGDYVSNIDK

```
SEQ2450        SVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
SEQ2451        SVQASEKVELKVATDSDTAPFTYXKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
SEQ2452        ----------------------------FKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
SEQ2453        SVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
SEQ2454        SVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
SEQ2455        SVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
SEQ2456        SVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
SEQ2457        SVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
SEQ2458        SVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
SEQ2459        SVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS
SEQ2460        SVQASEKVELKVATDSDTAPFTYQKDGKFKGYDVDVVKAVFKGSKYKVTFKTVPFDTIS

SEQ2450        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SEQ2451        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SEQ2452        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SEQ2453        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SEQ2454        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SEQ2455        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SEQ2456        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SEQ2457        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SEQ2458        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SEQ2459        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL
SEQ2460        TGIDAGKFDLSANDFSYNKERAEKYLFSDPISRSNYAVVGKKGSHYKSLSDLSGKSTEVL

SEQ2450        SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
SEQ2451        SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
SEQ2452        SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
SEQ2453        SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
SEQ2454        SGVNYAQVLENWNKNHXNKKPXKXKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
SEQ2455        SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
SEQ2456        SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
SEQ2457        SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
SEQ2458        SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
SEQ2459        SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK
SEQ2460        SGVNYAQVLENWNKNHPNKKPIKIKYVSGTTGVTSRLKNIESGKIDFILYDAISSDYIVK

SEQ2450        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKEDGTLARLSKQY
SEQ2451        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKENGTLARLSKQY
SEQ2452        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKK----------------------------
SEQ2453        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKENGTLARLSKQY
SEQ2454        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKR-----------------
SEQ2455        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKENGTLARLSKQY
SEQ2456        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKEDGTLARLSKQY
SEQ2457        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKENGTLARLSKQY
SEQ2458        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRNKVLKENG---------
SEQ2459        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKEDGTLARLSKQY
SEQ2460        DQSLNLSVSPLKGKIGNNKDGLEYLLLPKDKKGKTLQKFINKRIKVLKEDGTLARLSKQY
```

Figure 24: Comparative Sequences relating to SAG0290 (ABC transporter, substrate-binding protein)

SEQ2450	GGDYVSNIDK
SEQ2451	GGDYVSNIDK
SEQ2452	----------
SEQ2453	GGDYVSNIDK
SEQ2454	----------
SEQ2455	GGDYVSNIDK
SEQ2456	GGDYVSNIDK
SEQ2457	GGDYVSNIDK
SEQ2458	----------
SEQ2459	GGDYVSNIDK
SEQ2460	GGDYVSNIDK

Table 25: Comparative Sequences relating to SAG0368
(protein of unknown function)

SEQ ID NO. 2501: SAG0368 FROM THE 090 GBS TYPE Ia STRAIN
TATAATTTTTCGACTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCTATTGAA
GAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAAT
AGCGATTCTATGATCTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTG
ATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGAGTAGAAGCAAAGCTAAATGCAGCCTATGCTTCTGGTGGT
GCGGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGA
TTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAAT
GAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATG
CGCTATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATA
TTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGCAGTAAGTAATAACATGCAAACTAATATTGAGATA
TCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAA
GACGCTACTTTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAG
AAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACTACTGCT
AGTAATGATTCTTCTACTTATTCATCAACACAAGAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGT
TACAGTGGTAATACTACTTATAGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTCCTGCT
AGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCAATAATCATAACGGGGCTGCAACG
CCTAATCCA

SEQ ID NO. 2502: SAG0368 FROM THE 1169NT1 GBS TYPE V STRAIN
TATAATTTTTCGACTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCTATTGAA
GAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTTGGTCAGGAAA
TAGCGATTCTATGATCTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATT
GATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGCGTAGAAGCAAAGCTAAATGCAGCCTATGCTTCTGGTGG
TGCGGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGG
ATTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAA
TGAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTAT
GCGCTATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAAT
ATTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGCAGTAAGTAATAACATGCAAACTAATATTGAGAT
ATCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAAGGTGA
AGACGCTACTTTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAA
GAAAGAACTAGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACTACTGC
TAGTAATGATTCTTCTACTTATTCATCAACACAAGAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAG
TTACAGTGGTAATACTACTTATAGTTCTGAGACTAATCAAACAACTCATCAAAGTTACTATAATAGTAGCACTCCTGC
TAATAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCAATAATCATAATGGGGCTGCAAC
GCCTAATCCA

SEQ ID NO. 2503 SAG0368 FROM THE 18RS21 GBS TYPE II STRAIN
TATAATTTTTCGACTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCTATTGAA
GAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAAT
AGCGATTCTATGATCTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTG
ATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGAGTAGAAGCAAAGCTAAATGCAGCCTATGCTTCTGGTGGT
GCGGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGA
TTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAAT
GAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATG
CGCTATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATA
TTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGCAGTAAGTAATAACATGCAAACTAATATTGAGATA
TCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAA
GACGCTACTTTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAG
AAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACTACTGCT
AGTAATGATTCTTCTACTTATTCATCAACACAAGAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGT
TACAGTGGTAATACTACTTATAGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTCCTGCT
AGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCAATAATCATAACGGGGCTGCAACG
CCTAATCCA

Table 25: Comparative Sequences relating to SAG0368
(protein of unknown function)

SEQ ID NO. 2504: SAG0368 FROM THE 2603 V/R GBS TYPE V STRAIN

TATAATTTTTCGACTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCTATTGAA
GAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAAT
AGCGATTCTATGATCTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTG
ATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGAGTAGAAGCAAAGCTAAATGCAGCCTATGCTTCTGGTGGT
GCGGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGA
TTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAAT
GAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATG
CGCTATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATA
TTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGCAGTAAGTAATAACATGCAAACTAATATTGAGATA
TCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAA
GACGCTACTTTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAG
AAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACTACTGCT
AGTAATGATTCTTCTACTTATTCATCAACACAAGAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGT
TACAGTGGTAATACTACTTATAGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTCCTGCT
AGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCAATAATCATAACGGGGCTGCAACG
CCTAATCCA

SEQ ID NO. 2505: SAG0368 FROM THE A909 GBS TYPE Ia STRAIN

TATAATTTTTCGACTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCTATTGAA
GAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAAT
AGCGATTCTATGATCTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTG
ATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGAGTAGAAGCAAAGCTAAATGCAGCCTATGCTTCTGGTGGT
GCGGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGA
TTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAAT
GAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATG
CGCTATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATA
TTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGCAGTAAGTAATAACATGCAAACTAATATTGAGATA
TCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAA
GACGCTACTTTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAG
AAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACTACTGCT
AGTAATGATTCTTCTACTTATTCATCAACACAAGAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGT
TACAGTGGTAATACTACTTATAGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTCCTGCT
AGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCAATAATCATAACGGGGCTGCAACG
CCTAATCCA

SEQ ID NO. 2506: SAG0368 FROM THE CJB110 GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)

TATAATTTTTCGACTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCTATTGAA
GAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAAT
AGCGATTCTATGATCTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTG
ATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGAGTAGAAGCAAAGCTAAATGCAGCCTATGCTTCTGGTGGT
GCGGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGA
TTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAAT
GAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATG
CGCTATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATA
TTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGCAGTAAGTAATAACATGCAAACTAATATTGAGATA
TCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAA
GACGCTACTTTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAG
AAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACTACTGCT
AGTAATGATTCTTCTACTTATTCATCAACACAAGAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGT
TACAGTGGTAATACTACTTATTAGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTCCTGC
TAGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCAATAATCATAACGGGGCTGCAAC
GCCTAATCCA

Table 25: Comparative Sequences relating to SAG0368
(protein of unknown function)

SEQ ID NO. 2507: SAG0368 FROM THE COH1 GBS TYPE III STRAIN (REVERSE COMPLEMENT)

GATTTTAAGCTAGATAAATCAAAAAGTCATGCTATTGAAGAAACAAAGCCGTTTTCAATACTATTAATGGGTGTGGAC
ACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAATAGCGATTCTATGATCTTAGTCACTATAAATCCTAAAACT
AATAAAACAACGATGACAAGCTTAGAACGTGACGTATTGATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGC
GTAGAAGCAAAGCTAAATGCAGCCTATGCTTCTGGTGGTGCGGAAATGGCATTGATGACTGTTCAAGACTTATTAGAT
ATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGATTAGTTGATTTGGTCAATGCTGTTGGTGGTATAACAGTA
ACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAATGAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACAT
AAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATGCGCTATGATGATCCAGAGGGAGATTATGGGCGTCAAAAA
AGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATATTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTT
TCCGCAGTAAGTAATAACATGCAAACTAATATTGAGATATCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGAT
TCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAAGACGCTACTCTATCAGATGGTGGCTCTTATCAAATTTTA
ACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAGAAAGAGCTGGATAAAAAGCGTAGTAAAACTCTGAAGACA
AGCGCGATTCTATATGAAGATTACTATGGTACTACTGCTAGTAATGATTCTTCTACTTATTCATCAACACAAGAGAAT
TATTATTATACAACACCCTTATTCAGAAGCACCACCAAGTTACAGTGGTAATACTACTTATAGTTCTGAGACTAATCA
AACAACTCATCAAAGTTACTATAATAGTAGCACTCCTGCTAGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGA
TTCAAGTGGAAGTGTTAATAATTATAACGGGGCTGCAACGCCTAATCCAAACACAGGAACGCAACCAGTACCAGGTCA
AACTAATCCA

SEQ ID NO. 2508: SAG0368 FROM THE H36b GBS TYPE Ib STRAIN

TATAATTTTTCGACTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCTATTGAA
GAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAAT
AGCGATTCTATGATCTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTG
ATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGAGTAGAAGCAAAGCTAAATGCAGCCTATGCTTCTGGTGGT
GCGGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGA
TTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAAT
GAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATG
CGCTATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATA
TTGGCGTTAAATAGTA

SEQ ID NO. 2509: SAG0368 FROM THE __________

TTAGTTCATACAAAAAAATTCTTTCCGCAGTAAGTAATAACATGCAAACTAATATTGAGATATCATCAAAAACGATTC
CTAATTTGTTAGCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAAGACGCTACTTTATCAG
ATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAGAAAGAACTGGATAAAA
AGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACTACTGCTAGTAATGATTCTTCTA
CTTATTCATCAACACAAGAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGTTACAGTGGTAATACTA
CTTATAGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTCCTGCTAGTAACTATAGCAGTA
ACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCAATAATCATAACGGGGCTGCAACGCCTAATCCA

SEQ ID NO. 2510: SAG0368 FROM THE JM9130013 GBS TYPE VIII STRAIN (REVERSE COMPLEMENT)

TATAATTTTTCGACTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCTATTGAA
GAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAAT
AGCGATTCTATGATCTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTG
ATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGAGTAGAAGCAAAGCTAAATGCAGCCTATGCTTCTGGTGGT
GCGGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGA
TTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAAT
GAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACATAAAATAAATGGAGAACAAgCACTTGTTTATTCTCGTATG
CGCTATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATA
TTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGCAGTAAGTAATAACATGCAAACTAATATTGAGATA
TCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAA
GACGCTACTTTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAG
AAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACTACTGCT
AGTAATGATTCTTCTACTTATTCATCAACACAAGAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGT
TACAGTGGTAATACTACTTATAGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTCCTGCT
AGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCAATAATCATAACGGGGCTGCAACG
CCTAATCCA

Table 25: Comparative Sequences relating to SAG0368 (protein of unknown function)

SEQ ID NO. 2511: SAG0368 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)

```
TTCAATACTATTAATGGGTGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAATAGCGATTCTATGAT
CTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTGATTAAATTGAGTGG
TCCCAAAAATAATGGACAGACTGGCGTAGAAGCAAAGCTAAATGCAGCCTATGCTTCTGGTGGTGCGGAAATGGCATT
GATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGATTAGTTGATTTGGT
CAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAATGAACCAGAGTACAA
GGCTGTTGTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATGCGCTATGATGATCC
AGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATATTGGCGTTAAATAG
TATTAGTTCATACAAAAAAATTCTTTCCGCAGTAAGTAATAACATGCAAACTAATATTGAGATATCATCAAAAACGAT
TCCTAATTTGTTAGCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAAGACGCTACTCTATC
AGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAGAAAGAGCTGGATAA
AAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACTACTGCTAGTAATGATTCTTC
TACTTATTCATCAACACAAGAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGTTACAGTGGTAATAC
TACTTATAGTTCTGAGACTAATCAAACAACTCATCAAAGTTACTATAATAGTAGCACTCCTGCTAGTAACTATAGCAG
TAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTTAATAATTATAACGGGGCTGCAACGCCTAATCCAAACAC
AGGAACGCAACCAGTACCAGGTCAAACTAATCCA
```

```
SEQ2501             -----------------------------------------------------------
SEQ2502             -----------------------------------------------------------
SEQ2503             -----------------------------------------------------------
SEQ2504             -----------------------------------------------------------
SEQ2505             -----------------------------------------------------------
SEQ2506             -----------------------------------------------------------
SEQ2507             ATTTTAAGCTAGATAAATCAAAAAGTCATGCTATTGAAGAAACAAAGCCGTTTTCAATA
SEQ2508             -----------------------------------------------------------
SEQ2509             -----------------------------------------------------------
SEQ2510             -----------------------------------------------------------
SEQ2511             ----------------------------------------------------TTCAATA

SEQ2501             -----------------------------------------------------------
SEQ2502             -----------------------------------------------------------
SEQ2503             -----------------------------------------------------------
SEQ2504             -----------------------------------------------------------
SEQ2505             -----------------------------------------------------------
SEQ2506             -----------------------------------------------------------
SEQ2507             TATTAATGGGTGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAATAGC
SEQ2508             -----------------------------------------------------------
SEQ2509             -----------------------------------------------------------
SEQ2510             -----------------------------------------------------------
SEQ2511             TATTAATGGGTGTGGACACAGGTTCAGAGCATCGAAAATCTAAGTGGTCAGGAAATAGC

SEQ2501             -----------------------------------------------------------
SEQ2502             -----------------------------------------------------------
SEQ2503             -----------------------------------------------------------
SEQ2504             -----------------------------------------------------------
SEQ2505             -----------------------------------------------------------
SEQ2506             -----------------------------------------------------------
SEQ2507             ATTCTATGATCTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTA
SEQ2508             -----------------------------------------------------------
SEQ2509             -----------------------------------------------------------
SEQ2510             -----------------------------------------------------------
SEQ2511             ATTCTATGATCTTAGTCACTATAAATCCTAAAACTAATAAAACAACGATGACAAGCTTA
```

Table 25: Comparative Sequences relating to SAG0368 (protein of unknown function)

```
SEQ2501     -----------------------------------------------------------
SEQ2502     -----------------------------------------------------------
SEQ2503     -----------------------------------------------------------
SEQ2504     -----------------------------------------------------------
SEQ2505     -----------------------------------------------------------
SEQ2506     -----------------------------------------------------------
SEQ2507     AACGTGACGTATTGATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGCGTAGAA
SEQ2508     -----------------------------------------------------------
SEQ2509     -----------------------------------------------------------
SEQ2510     -----------------------------------------------------------
SEQ2511     AACGTGACGTATTGATTAAATTGAGTGGTCCCAAAAATAATGGACAGACTGGCGTAGAA

SEQ2501     -----------------------------------------------------------
SEQ2502     -----------------------------------------------------------
SEQ2503     -----------------------------------------------------------
SEQ2504     -----------------------------------------------------------
SEQ2505     -----------------------------------------------------------
SEQ2506     -----------------------------------------------------------
SEQ2507     CAAAGCTAAATGCAGCCTATGCTTCTGGTGGTGCGGAAATGGCATTGATGACTGTTCAA
SEQ2508     -----------------------------------------------------------
SEQ2509     -----------------------------------------------------------
SEQ2510     -----------------------------------------------------------
SEQ2511     CAAAGCTAAATGCAGCCTATGCTTCTGGTGGTGCGGAAATGGCATTGATGACTGTTCAA

SEQ2501     -----------------------------------------------------------
SEQ2502     -----------------------------------------------------------
SEQ2503     -----------------------------------------------------------
SEQ2504     -----------------------------------------------------------
SEQ2505     -----------------------------------------------------------
SEQ2506     -----------------------------------------------------------
SEQ2507     ACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGATTAGTTGAT
SEQ2508     -----------------------------------------------------------
SEQ2509     -----------------------------------------------------------
SEQ2510     -----------------------------------------------------------
SEQ2511     ACTTATTAGATATTAATGTTGATTACTTTATGCAAATTAATATGCAAGGATTAGTTGAT

SEQ2501     -----------------------------------------------------------
SEQ2502     -----------------------------------------------------------
SEQ2503     -----------------------------------------------------------
SEQ2504     -----------------------------------------------------------
SEQ2505     -----------------------------------------------------------
SEQ2506     -----------------------------------------------------------
SEQ2507     TGGTCAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATT
SEQ2508     -----------------------------------------------------------
SEQ2509     -----------------------------------------------------------
SEQ2510     -----------------------------------------------------------
SEQ2511     TGGTCAATGCTGTTGGTGGTATAACAGTAACTAATAAATTTGACTTTCCAATATCAATT

SEQ2501     -----------------------------------------------------------
SEQ2502     -----------------------------------------------------------
SEQ2503     -----------------------------------------------------------
SEQ2504     -----------------------------------------------------------
SEQ2505     -----------------------------------------------------------
SEQ2506     -----------------------------------------------------------
SEQ2507     CTGCCAATGAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACATAAAATAAATGGA
SEQ2508     -----------------------------------------------------------
SEQ2509     -----------------------------------------------------------
SEQ2510     -----------------------------------------------------------
SEQ2511     CTGCCAATGAACCAGAGTACAAGGCTGTTGTTGAACCAGGGACACATAAAATAAATGGA
```

Table 25: Comparative Sequences relating to SAG0368
(protein of unknown function)

```
SEQ2501        -----------------------------------------------------------
SEQ2502        -----------------------------------------------------------
SEQ2503        -----------------------------------------------------------
SEQ2504        -----------------------------------------------------------
SEQ2505        -----------------------------------------------------------
SEQ2506        -----------------------------------------------------------
SEQ2507        AACAAGCACTTGTTTATTCTCGTATGCGCTATGATGATCCAGAGGGAGATTATGGGCGT
SEQ2508        -----------------------------------------------------------
SEQ2509        -----------------------------------------------------------
SEQ2510        -----------------------------------------------------------
SEQ2511        AACAAGCACTTGTTTATTCTCGTATGCGCTATGATGATCCAGAGGGAGATTATGGGCGT

SEQ2501        -----------------------------------------------TATAATTTTTCG
SEQ2502        -----------------------------------------------TATAATTTTTCG
SEQ2503        -----------------------------------------------TATAATTTTTCG
SEQ2504        -----------------------------------------------TATAATTTTTCG
SEQ2505        -----------------------------------------------TATAATTTTTCG
SEQ2506        -----------------------------------------------TATAATTTTTCG
SEQ2507        AAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATATTGGCGTTAAATAGT
SEQ2508        -----------------------------------------------TATAATTTTTCG
SEQ2509        -----------------------------------------------------------
SEQ2510        -----------------------------------------------TATAATTTTTCG
SEQ2511        AAAAAAGACAACGTGAAGTAATTCAAAAAGTCCTTAAAAAAATATTGGCGTTAAATAGT

SEQ2501        CTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCT
SEQ2502        CTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCT
SEQ2503        CTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCT
SEQ2504        CTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCT
SEQ2505        CTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCT
SEQ2506        CTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCT
SEQ2507        TTAGTTCAT-ACAAAAAAATTCTTTCCGCAGTAAGTAA--TAACATGCAAACTAATATT
SEQ2508        CTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCT
SEQ2509        TTAGTTCAT-ACAAAAAAATTCTTTCCGCAGTAAGTAA--TAACATGCAAACTAATATT
SEQ2510        CTAATGAATTGTCTAAGACTTTTAAAGATTTTAAGCTAGCTAAATCAAAAAGTCATGCT
SEQ2511        TTAGTTCAT-ACAAAAAAATTCTTTCCGCAGTAAGTAA--TAACATGCAAACTAATATT

SEQ2501        TTGAAGAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCAT
SEQ2502        TTGAAGAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCAT
SEQ2503        TTGAAGAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCAT
SEQ2504        TTGAAGAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCAT
SEQ2505        TTGAAGAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCAT
SEQ2506        TTGAAGAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCAT
SEQ2507        AGATATCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGATTCA---TTGGAACAT
SEQ2508        TTGAAGAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCAT
SEQ2509        AGATATCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGATTCA---TTGGAACAT
SEQ2510        TTGAAGAAACAAAGCCGTTTTCAATACTATTAATGGGGGTGGACACAGGTTCAGAGCAT
SEQ2511        AGATATCATCAAAAACGATTCCTAATTTGTTAGCTTATAAAGATTCA---TTGGAACAT

SEQ2501        GAAAATCTAAGT-GGTCAGGAAATAGCGATTCTATGATCTTAGTCACTATAAATCCTAA
SEQ2502        GAAAATCTAAGTTGGTCAGGAAATAGCGATTCTATGATCTTAGTCACTATAAATCCTAA
SEQ2503        GAAAATCTAAGT-GGTCAGGAAATAGCGATTCTATGATCTTAGTCACTATAAATCCTAA
SEQ2504        GAAAATCTAAGT-GGTCAGGAAATAGCGATTCTATGATCTTAGTCACTATAAATCCTAA
SEQ2505        GAAAATCTAAGT-GGTCAGGAAATAGCGATTCTATGATCTTAGTCACTATAAATCCTAA
SEQ2506        GAAAATCTAAGT-GGTCAGGAAATAGCGATTCTATGATCTTAGTCACTATAAATCCTAA
SEQ2507        TTAAATCTTATC-AGTTGAAGGGTGAAGACGCTACTCTATCAG--ATGGTGGCTCTTAT
SEQ2508        GAAAATCTAAGT-GGTCAGGAAATAGCGATTCTATGATCTTAGTCACTATAAATCCTAA
SEQ2509        TTAAATCTTATC-AGTTGAAGGGTGAAGACGCTACTTTATCAG--ATGGTGGCTCTTAT
SEQ2510        GAAAATCTAAGT-GGTCAGGAAATAGCGATTCTATGATCTTAGTCACTATAAATCCTAA
SEQ2511        TTAAATCTTATC-AGTTGAAGGGTGAAGACGCTACTCTATCAG--ATGGTGGCTCTTAT

SEQ2501        ACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTGATTAAATTGAGTGGTCC
```

Table 25: Comparative Sequences relating to SAG0368 (protein of unknown function)

```
SEQ2502       ACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTGATTAAATTGAGTGGTCC
SEQ2503       ACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTGATTAAATTGAGTGGTCC
SEQ2504       ACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTGATTAAATTGAGTGGTCC
SEQ2505       ACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTGATTAAATTGAGTGGTCC
SEQ2506       ACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTGATTAAATTGAGTGGTCC
SEQ2507       AAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAGAAAGAGCTGGAT
SEQ2508       ACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTGATTAAATTGAGTGGTCC
SEQ2509       AAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAGAAAGAACTGGAT
SEQ2510       ACTAATAAAACAACGATGACAAGCTTAGAACGTGACGTATTGATTAAATTGAGTGGTCC
SEQ2511       AAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAATAGAATTAAGAAAGAGCTGGAT

SEQ2501       AAAAATAATGGACAGACTGGAGTAGAAGCAAAG--CTAAATGCAGCCTATGCTTCTGGT
SEQ2502       AAAAATAATGGACAGACTGGCGTAGAAGCAAAG--CTAAATGCAGCCTATGCTTCTGGT
SEQ2503       AAAAATAATGGACAGACTGGAGTAGAAGCAAAG--CTAAATGCAGCCTATGCTTCTGGT
SEQ2504       AAAAATAATGGACAGACTGGAGTAGAAGCAAAG--CTAAATGCAGCCTATGCTTCTGGT
SEQ2505       AAAAATAATGGACAGACTGGAGTAGAAGCAAAG--CTAAATGCAGCCTATGCTTCTGGT
SEQ2506       AAAAATAATGGACAGACTGGAGTAGAAGCAAAG--CTAAATGCAGCCTATGCTTCTGGT
SEQ2507       AAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACT
SEQ2508       AAAAATAATGGACAGACTGGAGTAGAAGCAAAG--CTAAATGCAGCCTATGCTTCTGGT
SEQ2509       AAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACT
SEQ2510       AAAAATAATGGACAGACTGGAGTAGAAGCAAAG--CTAAATGCAGCCTATGCTTCTGGT
SEQ2511       AAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCTATATGAAGATTACTATGGTACT

SEQ2501       GTGC-GGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTT
SEQ2502       GTGC-GGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTT
SEQ2503       GTGC-GGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTT
SEQ2504       GTGC-GGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTT
SEQ2505       GTGC-GGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTT
SEQ2506       GTGC-GGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTT
SEQ2507       CTGCTAGTAATGATTCTTCTACTTATTCATCAAC-ACAAGAGAATTATTATTAT-ACAA
SEQ2508       GTGC-GGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTT
SEQ2509       CTGCTAGTAATGATTCTTCTACTTATTCATCAAC-ACAAGAGAATAATTATAAT-ACAA
SEQ2510       GTGC-GGAAATGGCATTGATGACTGTTCAAGACTTATTAGATATTAATGTTGATTACTT
SEQ2511       CTGCTAGTAATGATTCTTCTACTTATTCATCAAC-ACAAGAGAATAATTATAAT-ACAA

SEQ2501       ATGCAAATTAATATGCAAGGATTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGT
SEQ2502       ATGCAAATTAATATGCAAGGATTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGT
SEQ2503       ATGCAAATTAATATGCAAGGATTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGT
SEQ2504       ATGCAAATTAATATGCAAGGATTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGT
SEQ2505       ATGCAAATTAATATGCAAGGATTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGT
SEQ2506       ATGCAAATTAATATGCAAGGATTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGT
SEQ2507       ACCCTTATTCAGAAGCACCACCAAGTTACAGTGGT-AATACTACTTATAGTT---CTGA
SEQ2508       ATGCAAATTAATATGCAAGGATTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGT
SEQ2509       ACC-TTATTCAGAAGCACCACCAAGTTACAGTGGT-AATACTACTTATAGTT---CTGA
SEQ2510       ATGCAAATTAATATGCAAGGATTAGTTGATTTAGTCAATGCTGTTGGTGGTATAACAGT
SEQ2511       ACC-TTATTCAGAAGCACCACCAAGTTACAGTGGT-AATACTACTTATAGTT---CTGA

SEQ2501       ACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAATGAACCAGAGTACAAGGCTGT
SEQ2502       ACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAATGAACCAGAGTACAAGGCTGT
SEQ2503       ACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAATGAACCAGAGTACAAGGCTGT
SEQ2504       ACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAATGAACCAGAGTACAAGGCTGT
SEQ2505       ACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAATGAACCAGAGTACAAGGCTGT
SEQ2506       ACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAATGAACCAGAGTACAAGGCTGT
SEQ2507       ACTAATCAAAC-AACTCATCAA----AGTTACTAT-AATAG--TAGCACTCCTGCTAGT
SEQ2508       ACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAATGAACCAGAGTACAAGGCTGT
SEQ2509       ACTAATCAAAC-AACTCATCAA----AATTACTAT-AATAG--TAGCACTCCTGCTAGT
SEQ2510       ACTAATAAATTTGACTTTCCAATATCAATTGCTGCCAATGAACCAGAGTACAAGGCTGT
SEQ2511       ACTAATCAAAC-AACTCATCAA----AGTTACTAT-AATAG--TAGCACTCCTGCTAGT
```

Table 25: Comparative Sequences relating to SAG0368 (protein of unknown function)

```
SEQ2501         GTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATGCG
SEQ2502         GTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATGCG
SEQ2503         GTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATGCG
SEQ2504         GTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATGCG
SEQ2505         GTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATGCG
SEQ2506         GTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATGCG
SEQ2507         ACTATAGCAGTAACAC-TAACACAGGTCAGGCTGATTCAAGTGGAAGTGTTAATAATTA
SEQ2508         GTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATGCG
SEQ2509         ACTATAGCAGTAACAC-TAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCAATAATCA
SEQ2510         GTTGAACCAGGGACACATAAAATAAATGGAGAACAAGCACTTGTTTATTCTCGTATGCG
SEQ2511         ACTATAGCAGTAACAC-TAACACAGGTCAGGCTGATTCAAGTGGAAGTGTTAATAATTA

SEQ2501         TATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAA
SEQ2502         TATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAA
SEQ2503         TATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAA
SEQ2504         TATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAA
SEQ2505         TATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAA
SEQ2506         TATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAA
SEQ2507         AACGGGGCTGCAACGCCTAATCCAAACACAGGAACGCAACCAGTACCAGGTCAAACTAA
SEQ2508         TATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAA
SEQ2509         AACGGGGCTGCAACGCCTAATCCA-----------------------------------
SEQ2510         TATGATGATCCAGAGGGAGATTATGGGCGTCAAAAAAGACAACGTGAAGTAATTCAAAA
SEQ2511         AACGGGGCTGCAACGCCTAATCCAAACACAGGAACGCAACCAGTACCAGGTCAAACTAA

SEQ2501         GTCCTTAAAAAAATATTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGC
SEQ2502         GTCCTTAAAAAAATATTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGC
SEQ2503         GTCCTTAAAAAAATATTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGC
SEQ2504         GTCCTTAAAAAAATATTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGC
SEQ2505         GTCCTTAAAAAAATATTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGC
SEQ2506         GTCCTTAAAAAAATATTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGC
SEQ2507         CCA--------------------------------------------------------
SEQ2508         GTCCTTAAAAAAATATTGGCGTTAAATAGTA----------------------------
SEQ2509         -----------------------------------------------------------
SEQ2510         GTCCTTAAAAAAATATTGGCGTTAAATAGTATTAGTTCATACAAAAAAATTCTTTCCGC
SEQ2511         CCA--------------------------------------------------------

SEQ2501         GTAAGTAATAACATGCAAACTAATATTGAGATATCATCAAAAACGATTCCTAATTTGTT
SEQ2502         GTAAGTAATAACATGCAAACTAATATTGAGATATCATCAAAAACGATTCCTAATTTGTT
SEQ2503         GTAAGTAATAACATGCAAACTAATATTGAGATATCATCAAAAACGATTCCTAATTTGTT
SEQ2504         GTAAGTAATAACATGCAAACTAATATTGAGATATCATCAAAAACGATTCCTAATTTGTT
SEQ2505         GTAAGTAATAACATGCAAACTAATATTGAGATATCATCAAAAACGATTCCTAATTTGTT
SEQ2506         GTAAGTAATAACATGCAAACTAATATTGAGATATCATCAAAAACGATTCCTAATTTGTT
SEQ2507         -----------------------------------------------------------
SEQ2508         -----------------------------------------------------------
SEQ2509         -----------------------------------------------------------
SEQ2510         GTAAGTAATAACATGCAAACTAATATTGAGATATCATCAAAAACGATTCCTAATTTGTT
SEQ2511         -----------------------------------------------------------

SEQ2501         GCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAAGACGCTAC
SEQ2502         GCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAAGGTGAAGACGCTAC
SEQ2503         GCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAAGACGCTAC
SEQ2504         GCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAAGACGCTAC
SEQ2505         GCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAAGACGCTAC
SEQ2506         GCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAAGACGCTAC
SEQ2507         -----------------------------------------------------------
SEQ2508         -----------------------------------------------------------
SEQ2509         -----------------------------------------------------------
SEQ2510         GCTTATAAAGATTCATTGGAACATATTAAATCTTATCAGTTGAAGGGTGAAGACGCTAC
SEQ2511         -----------------------------------------------------------
```

Table 25: Comparative Sequences relating to SAG0368 (protein of unknown function)

```
SEQ2501        TTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAA
SEQ2502        TTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAA
SEQ2503        TTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAA
SEQ2504        TTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAA
SEQ2505        TTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAA
SEQ2506        TTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAA
SEQ2507        -----------------------------------------------------------
SEQ2508        -----------------------------------------------------------
SEQ2509        -----------------------------------------------------------
SEQ2510        TTATCAGATGGTGGCTCTTATCAAATTTTAACTAAGAAACATCTACTTGCAGTTCAAAA
SEQ2511        -----------------------------------------------------------

SEQ2501        AGAATTAAGAAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCT
SEQ2502        AGAATTAAGAAAGAACTAGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCT
SEQ2503        AGAATTAAGAAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCT
SEQ2504        AGAATTAAGAAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCT
SEQ2505        AGAATTAAGAAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCT
SEQ2506        AGAATTAAGAAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCT
SEQ2507        -----------------------------------------------------------
SEQ2508        -----------------------------------------------------------
SEQ2509        -----------------------------------------------------------
SEQ2510        AGAATTAAGAAAGAACTGGATAAAAAGCGTAGTAAAACTCTGAAGACAAGCGCGATTCT
SEQ2511        -----------------------------------------------------------

SEQ2501        TATGAAGATTACTATGGTACTACTGCTAGTAATGATTCTTCTACTTATTCATCAACACA
SEQ2502        TATGAAGATTACTATGGTACTACTGCTAGTAATGATTCTTCTACTTATTCATCAACACA
SEQ2503        TATGAAGATTACTATGGTACTACTGCTAGTAATGATTCTTCTACTTATTCATCAACACA
SEQ2504        TATGAAGATTACTATGGTACTACTGCTAGTAATGATTCTTCTACTTATTCATCAACACA
SEQ2505        TATGAAGATTACTATGGTACTACTGCTAGTAATGATTCTTCTACTTATTCATCAACACA
SEQ2506        TATGAAGATTACTATGGTACTACTGCTAGTAATGATTCTTCTACTTATTCATCAACACA
SEQ2507        -----------------------------------------------------------
SEQ2508        -----------------------------------------------------------
SEQ2509        -----------------------------------------------------------
SEQ2510        TATGAAGATTACTATGGTACTACTGCTAGTAATGATTCTTCTACTTATTCATCAACACA
SEQ2511        -----------------------------------------------------------

SEQ2501        GAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGTTACAGTGGTAATAC
SEQ2502        GAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGTTACAGTGGTAATAC
SEQ2503        GAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGTTACAGTGGTAATAC
SEQ2504        GAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGTTACAGTGGTAATAC
SEQ2505        GAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGTTACAGTGGTAATAC
SEQ2506        GAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGTTACAGTGGTAATAC
SEQ2507        -----------------------------------------------------------
SEQ2508        -----------------------------------------------------------
SEQ2509        -----------------------------------------------------------
SEQ2510        GAGAATAATTATAATACAACACCTTATTCAGAAGCACCACCAAGTTACAGTGGTAATAC
SEQ2511        -----------------------------------------------------------

SEQ2501        ACTTAT-AGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTC
SEQ2502        ACTTAT-AGTTCTGAGACTAATCAAACAACTCATCAAAGTTACTATAATAGTAGCACTC
SEQ2503        ACTTAT-AGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTC
SEQ2504        ACTTAT-AGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTC
SEQ2505        ACTTAT-AGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTC
SEQ2506        ACTTATTAGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTC
SEQ2507        -----------------------------------------------------------
SEQ2508        -----------------------------------------------------------
SEQ2509        -----------------------------------------------------------
SEQ2510        ACTTAT-AGTTCTGAGACTAATCAAACAACTCATCAAAATTACTATAATAGTAGCACTC
SEQ2511        -----------------------------------------------------------
```

Table 25: Comparative Sequences relating to SAG0368 (protein of unknown function)

SEQ2501	TGCTAGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCA
SEQ2502	TGCTAATAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCA
SEQ2503	TGCTAGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCA
SEQ2504	TGCTAGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCA
SEQ2505	TGCTAGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCA
SEQ2506	TGCTAGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCA
SEQ2507	---
SEQ2508	---
SEQ2509	---
SEQ2510	TGCTAGTAACTATAGCAGTAACACTAACACAGGTCAGGCTGATTCAAGTGGAAGTGTCA
SEQ2511	---

SEQ2501	TAATCATAACGGGGCTGCAACGCCTAATCCA
SEQ2502	TAATCATAATGGGGCTGCAACGCCTAATCCA
SEQ2503	TAATCATAACGGGGCTGCAACGCCTAATCCA
SEQ2504	TAATCATAACGGGGCTGCAACGCCTAATCCA
SEQ2505	TAATCATAACGGGGCTGCAACGCCTAATCCA
SEQ2506	TAATCATAACGGGGCTGCAACGCCTAATCCA
SEQ2507	-------------------------------
SEQ2508	-------------------------------
SEQ2509	-------------------------------
SEQ2510	TAATCATAACGGGGCTGCAACGCCTAATCCA
SEQ2511	-------------------------------

>SEQ ID NO 2550: 54_090 frame: 1
YNFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
INPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
DYFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
RMRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
NLLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
AILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
STPASNYSSNTNTGQADSSGSVNNHNGAATPNP

>SEQ ID NO 2551:54_1169NT frame: 1
YNFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKLVRK.RFYDLSH
YKS.N..NNDDKLRT.RID.IEWSQK.WTDWRRSKAKCSLCFWWCGNGIDDCSRLIRY.C
.LLYAN.YARIS.FSQCCWWYNSN..I.LSNINCCQ.TRVQGCC.TRDT.NKWRTSTCLF
SYAL..SRGRLWASKKTT.SNSKSP.KNIGVK.Y.FIQKNSFRSK..HAN.Y.DIIKNDS
.FVSL.RFIGTY.ILSVER.RRYFIRWWLLSNFN.ETSTCSSK.N.ERTR.KA..NSEDK
RDSI.RLLWYYC...FFYLFINTRE.L.YNTLFRSTTKLQW.YYL.F.D.SNNSSKLL..
.HSC..L.Q.H.HRSG.FKWKCQ.S.WGCNA.S

>SEQ ID NO 2552:54_18RS21 frame: 1
YNFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
INPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
DYFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
RMRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
NLLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
AILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
STPASNYSSNTNTGQADSSGSVNNHNGAATPNP

>SEQ ID NO 2553:54_2603 frame: 1
YNFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
INPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
DYFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
RMRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
NLLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
AILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
STPASNYSSNTNTGQADSSGSVNNHNGAATPNP

>SEQ ID NO 2554: 54_A909 frame: 1
YNFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
INPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
DYFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
RMRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
NLLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS

Table 25: Comparative Sequences relating to SAG0368 (protein of unknown function)

AILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
STPASNYSSNTNTGQADSSGSVNNHNGAATPNP

>SEQ ID NO 2555:54_CJB110 frame: 1
YNFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
INPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
DYFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
RMRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
NLLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
AILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTY.F.D.SNNSSKLL..

>SEQ ID NO 2556:54_COH1 frame: 1
DFKLDKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVTINPKTNKTTMTSL
ERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINVDYFMQINMQGLVD
LVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYSRMRYDDPEGDYGR
QKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIPNLLAYKDSLEHIK
SYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTSAILYEDYYGTTAS
NDSSTYSSTQENYYYTTPLFRSTTKLQW.YYL.F.D.SNNSSKLL...HSC..L.Q.H.H
RSG.FKWKC..L.RGCNA.SKHRNATSTRSN.S

>SEQ ID NO 2557:54_H36B frame: 1
YNFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
INPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
DYFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
RMRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
NLLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
AILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
STPASNYSSNTNTGQADSSGSVNNHNGAATPNP

>SEQ ID NO 2558:54_JM9130013 frame: 1
YNFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
INPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
DYFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
RMRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
NLLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
AILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
STPASNYSSNTNTGQADSSGSVNNHNGAATPNP

>SEQ ID NO 2559:54_M781 frame: 2
SILLMGVDTGSEHRKSKWSGNSDSMILVTINPKTNKTTMTSLERDVLIKLSGPKNNGQTG
VEAKLNAAYASGGAEMALMTVQDLLDINVDYFMQINMQGLVDLVNAVGGITVTNKFDFPI
SIAANEPEYKAVVEPGTHKINGEQALVYSRMRYDDPEGDYGRQKRQREVIQKVLKKILAL
NSISSYKKILSAVSNNMQTNIEISSKTIPNLLAYKDSLEHIKSYQLKGEDATLSDGGSYQ
ILTKKHLLAVQNRIKKELDKKRSKTLKTSAILYEDYYGTTASNDSSTYSSTQENNYNTTP
YSEAPPSYSGNTTYSSETNQTTHQSYYNSSTPASNYSSNTNTGQADSSGSVNNYNGAATP
NPNTGTQPVPGQTNP

```
SEQ2550        NFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
SEQ2551        NFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKLVRKRFYDLSHY
SEQ2552        NFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
SEQ2553        NFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
SEQ2554        NFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
SEQ2555        NFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
SEQ2556        ------------DFKLDKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
SEQ2557        NFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
SEQ2558        NFSTNELSKTFKDFKLAKSKSHAIEETKPFSILLMGVDTGSEHRKSKWSGNSDSMILVT
SEQ2559        ------------------------------SILLMGVDTGSEHRKSKWSGNSDSMILVT

SEQ2550        NPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
SEQ2551        SNNNDDKLRTRIDIEWSQKWTDWRRS-----KAKCSLCFWWCGNGIDDCSRLIRYCLLY
SEQ2552        NPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
SEQ2553        NPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
SEQ2554        NPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
SEQ2555        NPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
SEQ2556        NPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
SEQ2557        NPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
```

Table 25: Comparative Sequences relating to SAG0368 (protein of unknown function)

```
SEQ2558         NPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV
SEQ2559         NPKTNKTTMTSLERDVLIKLSGPKNNGQTGVEAKLNAAYASGGAEMALMTVQDLLDINV

SEQ2550         YFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
SEQ2551         NYARISFSQCCWWYNS----------NILSNINCCQTRVQGCCTRDTNKWRTSTCLFSY
SEQ2552         YFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
SEQ2553         YFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
SEQ2554         YFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
SEQ2555         YFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
SEQ2556         YFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
SEQ2557         YFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
SEQ2558         YFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS
SEQ2559         YFMQINMQGLVDLVNAVGGITVTNKFDFPISIAANEPEYKAVVEPGTHKINGEQALVYS

SEQ2550         MRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
SEQ2551         LSRGRLWASKKTTSNSKSPKNIGVKYFIQKNSFRSKHANYDIIKNDSFVSLRFIGTYI-
SEQ2552         MRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
SEQ2553         MRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
SEQ2554         MRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
SEQ2555         MRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
SEQ2556         MRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
SEQ2557         MRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
SEQ2558         MRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP
SEQ2559         MRYDDPEGDYGRQKRQREVIQKVLKKILALNSISSYKKILSAVSNNMQTNIEISSKTIP

SEQ2550         LLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
SEQ2551         L-SVERRRYFIRWWLLSNFNETSTCSSKNERTRKANSEDKRDSIRLLWYYCFFYLFINT
SEQ2552         LLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
SEQ2553         LLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
SEQ2554         LLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
SEQ2555         LLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
SEQ2556         LLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
SEQ2557         LLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
SEQ2558         LLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS
SEQ2559         LLAYKDSLEHIKSYQLKGEDATLSDGGSYQILTKKHLLAVQNRIKKELDKKRSKTLKTS

SEQ2550         ILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
SEQ2551         ELY-----------NTLFRST--------------TKLQWYYLFDSNNSSKLLHSCLQH
SEQ2552         ILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
SEQ2553         ILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
SEQ2554         ILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
SEQ2555         ILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYFDSNNSSKLL-----
SEQ2556         ILYEDYYGTTASNDSSTYSSTQENYYYTTPLFRSTTKLQWYYLFDSNNSSKLLHSCLQH
SEQ2557         ILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
SEQ2558         ILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQNYYNS
SEQ2559         ILYEDYYGTTASNDSSTYSSTQENNYNTTPYSEAPPSYSGNTTYSSETNQTTHQSYYNS

SEQ2550         TPASNYSSNTNTGQADSSGSVNNHNGAATPNP-------------
SEQ2551         RSGFKWKCQSWGCNAS-----------------------------
SEQ2552         TPASNYSSNTNTGQADSSGSVNNHNGAATPNP-------------
SEQ2553         TPASNYSSNTNTGQADSSGSVNNHNGAATPNP-------------
SEQ2554         TPASNYSSNTNTGQADSSGSVNNHNGAATPNP-------------
SEQ2555         ---------------------------------------------
SEQ2556         RSGFKWKCLRGCNASKHRNATSTRSNS------------------
SEQ2557         TPASNYSSNTNTGQADSSGSVNNHNGAATPNP-------------
SEQ2558         TPASNYSSNTNTGQADSSGSVNNHNGAATPNP-------------
SEQ2559         TPASNYSSNTNTGQADSSGSVNNYNGAATPNPNTGTQPVPGQTNP
```

Table 26: Comparative Sequences relating to SAG0503 (lipase/acylhydolase)

SEQ ID NO. 2601: SAG0503 FROM THE 090 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)

GGGCACAAGTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAATCCTAAATTAACAAAAAA
AGACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGCTCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATAC
AACCTCTCAAGGTGGTTTTGTCCCACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAA
TTATGGTGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGAAAAAGATTTAGA
GAAAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGCTGTTATTCGTAAAGAGCTCAGTCATTTATC
ACTAAATTCCTTTGAGAAACCAGCAGAAGCATATAAGGAACGTTTGAAAGAAATACTTGCAAAAGCAAGACAAGATAA
TCCTAAATTGCCTATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAATGCAAAC
CGTTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAATGTTTATTTTGTCCCAATTAATGA
CCGCCTTTATAAGGGAATAAATGGTAAAGAGGGTATTACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGC
TCTCTTTACTGGAGACCATTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAA
TGAAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAAG

SEQ ID NO. 2602: SAG0503 FROM THE H36b GBS TYPE Ib STRAIN (REVERSE COMPLEMENT)

TTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAATCCTAAATTAACAAAAAAAGACTTCCT
AACAAAGAAAGTTATCCCACTTAACTATGTTGCTCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCTCA
AGGTGGTTTTGTTCCACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGGTGT
GTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGAAAAAGATTTAGAGAAAGCTGA
TTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGCTGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATTC
CTTTGAGAAACCAGCAGAAGCATATAAGGAACGTTTGAAAGAAATCCTTGCAAAAGCAAGACAAGATAATCCTAAATT
GCCTATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAATGCAAACCGTTATTGA
TAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAATGTTTATTTTGTCCCAATTAATGACCGCCTTTA
TAAGGGAATAAATGGTAAAGAGGGTATTATAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTAC
TGGAGACCATTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAATGAAACAAG
AAAAAACTGGCCGAACCCAGCTTTCTTGTACAAAGTGGTCC

SEQ ID NO. 2603: SAG0503 FROM THE 18RS21 GBS TYPE II STRAIN (REVERSE COMPLEMENT)

GTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAATCCTAAATTAACAAAAAAAGACTTCC
TAACAAAGAAAGTTATCCCACTTAACTATGTTGCTCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCTC
AAGGTGGTTTTGTTCCACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGGTG
TGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGAAAAAGATTTAGAGAAAGCTG
ATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGCTGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATT
CCTTTGAGAAACCAGCAGAAGCATATAAGGAACGTTTGAAAGAAATCCTTGCAAAAGCAAGACAAGATAATCCTAAAT
TGCCTATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAATGCAAACCGTTATTG
ATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAATGTTTATTTTGTCCCAATTAATGACCGCCTTT
ATAAGGGAATAAATGGTAAAGAGGGTATTACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTA
CTGGAGACCATTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAATGAAACAA
GAAAAAACTGGCCGAACCCAGCTTTCTTGTACAA

SEQ ID NO. 2604: SAG0503 FROM THE COH1 GBS TYPE III STRAIN (REVERSE COMPLEMENT)

GGACAAGTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAATCCTAAATTAACAAAAAAAG
ACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGCTCTTGGAGATTCTCTGACCGAAGGTGTGGGGGATACAA
CCTCTCAAGGTGGTTTTGTCCCACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATT
ATGGTGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGAAAAAGATTTAGAGA
AAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGCTGTTATTCGTAAAGAGCTCAGTCATTTATCAC
TAAATTCCTTTGAGAAACCAGCAGAAGCATATAAGGAACGTTTGAAAGAAATTCTTGCAAAAGCAAGACAAGATAATC
CTAAATTGCCTATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAATGCAAACCG
TTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAATGTTTATTTTGTCCCAATTAATGACC
GCCTTTATAAGGGAATAAATGGTAAAGAGGGTATTACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTC
TCTTTACTGGAGACCATTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAATG
AAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAA

Table 26: Comparative Sequences relating to SAG0503 (lipase/acylhydolase)

SEQ ID NO. 2605: SAG0503 FROM THE CJB110 GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)

GTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAATCCTAAATTAACAAAAAAAGACTTCC
TAACAAAGAAAGTTATCCCACTTAACTATGTTGCTCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCTC
AAGGTGGTTTTGTCCCACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGGTG
TGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGAAAAAGATTTAGAGAAAGCTG
ATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGCTGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATT
CCTTTGAGAAACCAGCAGAAGCATATAAGGAACGTTTGAAAGAAATACTTGCAAAAGCAAGACAAGATAATCCTAAAT
TGCCTATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAATGCAAACCGTTATTG
ATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAATGTTTATTTTGTCCCAATTAATGACCGCCTTT
ATAAGGGAATAAATGGTAAAGAGGGTATTACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTA
CTGGAGACCATTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAATGAAACAA
GAAAAAACTGGCCGAACCCAGCTTTCTTGTACAA

SEQ ID NO. 2606: SAG0503 FROM THE 1169NT1 GBS TYPE V STRAIN (REVERSE COMPLEMENT)

GTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAATCCTAAATTAACAAAAAAAGACTTCC
TAACAAAGAAAGTTATCCCACTTAACTATGTTGCTCTTGGAGATTCTCTGACCGAAGGTGTGGGGGATACAACCTCTC
AAGGTGGTTTTGTCCCACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGGTG
TGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGAAAAAGATTTAGAGAAAGCTG
ATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGCTGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATT
CCTTTGAGAAACCAGCAGAAGCATATAAGGAACGTTTGAAAGAAATTCTTGCAAAAGCAAGACAAGATAATCCTAAAT
TGCCTATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAATGCAAACCGTTATTG
ATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAATGTTTATTTTGTCCCAATTAATGACCGCCTTT
ATAAGGGAATAAATGGTAAAGAGGGTATTACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTA
CTGGAGACCATTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAATGAAACAA
GAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAA

SEQ ID NO. 2607: SAG0503 FROM THE JM9130013 GBS TYPE VIII STRAIN (REVERSE COMPLEMENT)

GTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAATCCTAAATTAACAAAAAAAGACTTCC
TAACAAAGAAAGTTATCCCACTTAACTATGTTGCTCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCTC
AAGGTGGTTTTGTTCCACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGGTG
TGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGAAAAAGATTTAGAGAAAGCTG
ATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGCTGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATT
CCTTTGAGAAACCAGCAGAAGCATATAAGGAACGTTTGAAAGAAATCCTTGCAAAAGCAAGACAAGATAATCCTAAAT
TGCCTATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAATGCAAACCGTTATTG
ATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAATGTTTATTTTGTCCCAATTAATGACCGCCTTT
ATAAGGGAATAAATGGTAAAGAGGGTATTACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTA
CTGGAGACCATTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAATGAAACAA
GAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAA

SEQ ID NO. 2608: SAG0503 FROM THE 2603 V/R GBS TYPE V STRAIN (REVERSE COMPLEMENT)

AGTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAATCCTAAATTAACAAAAAAAGACTTC
CTAACAAAGAAAGTTATCCCACTTAACTATGTTGCTCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCT
CAAGGTGGTTTTGTTCCACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGGT
GTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGAAAAAGATTTAGAGAAAGCT
GATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGCTGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAAT
TCCTTTGAGAAACCAGCAGAAGCATATAAGGAACGTTTGAAAGAAATCCTTGCAAAAGCAAGACAAGATAATCCTAAA
TTGCCTATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAATGCAAACCGTTATT
GATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAATGTTTATTTTGTCCCAATTAATGACCGCCTT
TATAAGGGAATAAATGGTAAAGAGGGTATTACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTT
ACTGGAGACCATTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAATGAAACA
AGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAAGTGG

Table 26: Comparative Sequences relating to SAG0503 (lipase/acylhydolase)

SEQ ID NO. 2609: SAG0503 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)

GGACAAGTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAATCCTAAATTAACAAAAAAAG
ACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGCTCTTGGAGATTCTCTGACCGAAGGTGTGGGGGATACAA
CCTCTCAAGGTGGTTTTGTCCCACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATT
ATGGTGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGAAAAAGATTTAGAGA
AAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGCTGTTATTCGTAAAGAGCTCAGTCATTTATCAC
TAAATTCCTTTGAGAAACCAGCAGAAGCATATAAGGAACGTTTGAAAGAAATTCTTGCAAAAGCAAGACAAGATAATC
CTAAATTGCCTATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAATGCAAACCG
TTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAATGTTTATTTTGTCCCAATTAATGACC
GCCTTTATAAGGGAATAAATGGTAAAGAGGGTATTACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTC
TCTTTACTGGAGACCATTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAATG
AAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAA

```
SEQ2601        GGCACAAGTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAA
SEQ2602        --------TTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAA
SEQ2603        -------      GTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAA
SEQ2604          GGACAAGTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAA
SEQ2605        -------GTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAA
SEQ2606        -------GTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAA
SEQ2607        -------GTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAA
SEQ2608        ------AGTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAA
SEQ2609          GGACAAGTTTGTACAAAAAAGCAGGCTCTATTTTTTCCTTGATCATTCCAAAATCAAA

SEQ2601        TCCTAAATTAACAAAAAAAGACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGC
SEQ2602        TCCTAAATTAACAAAAAAAGACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGC
SEQ2603        TCCTAAATTAACAAAAAAAGACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGC
SEQ2604        TCCTAAATTAACAAAAAAAGACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGC
SEQ2605        TCCTAAATTAACAAAAAAAGACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGC
SEQ2606        TCCTAAATTAACAAAAAAAGACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGC
SEQ2607        TCCTAAATTAACAAAAAAAGACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGC
SEQ2608        TCCTAAATTAACAAAAAAAGACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGC
SEQ2609        TCCTAAATTAACAAAAAAAGACTTCCTAACAAAGAAAGTTATCCCACTTAACTATGTTGC

SEQ2601        TCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCTCAAGGTGGTTTTGTCCC
SEQ2602        TCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCTCAAGGTGGTTTTGTTCC
SEQ2603        TCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCTCAAGGTGGTTTTGTTCC
SEQ2604        TCTTGGAGATTCTCTGACCGAAGGTGTGGGGGATACAACCTCTCAAGGTGGTTTTGTCCC
SEQ2605        TCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCTCAAGGTGGTTTTGTCCC
SEQ2606        TCTTGGAGATTCTCTGACCGAAGGTGTGGGGGATACAACCTCTCAAGGTGGTTTTGTCCC
SEQ2607        TCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCTCAAGGTGGTTTTGTTCC
SEQ2608        TCTTGGAGATTCTCTGACCGAAGGTGTGGGCGATACAACCTCTCAAGGTGGTTTTGTTCC
SEQ2609        TCTTGGAGATTCTCTGACCGAAGGTGTGGGGGATACAACCTCTCAAGGTGGTTTTGTCCC

SEQ2601        ACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGG
SEQ2602        ACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGG
SEQ2603        ACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGG
SEQ2604        ACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGG
SEQ2605        ACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGG
SEQ2606        ACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGG
SEQ2607        ACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGG
SEQ2608        ACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGG
SEQ2609        ACTGCTATCAGAATCACTCCATAATCGATACTCTTACCAAGTGACTTCTGTTAATTATGG

SEQ2601        TGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGA
SEQ2602        TGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGA
SEQ2603        TGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGA
SEQ2604        TGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGA
SEQ2605        TGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGA
SEQ2606        TGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGA
SEQ2607        TGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGA
SEQ2608        TGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGA
SEQ2609        TGTGTCTGGGAATACTAGTCAACAAATTTTAAAACGTATGACGACAGATCCTCAAATCGA

SEQ2601        AAAAGATTTAGAGAAAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGC
SEQ2602        AAAAGATTTAGAGAAAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGC
```

Table 26: Comparative Sequences relating to SAG0503 (lipase/acylhydolase)

SEQ2603 AAAAGATTTAGAGAAAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGC

SEQ2604 AAAAGATTTAGAGAAAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGC
SEQ2605 AAAAGATTTAGAGAAAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGC
SEQ2606 AAAAGATTTAGAGAAAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGC
SEQ2607 AAAAGATTTAGAGAAAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGC
SEQ2608 AAAAGATTTAGAGAAAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGC
SEQ2609 AAAAGATTTAGAGAAAGCTGATTTATTGACGCTAACTGTTGGTGGTAATGATGTCTTGGC

SEQ2601 TGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATTCCTTTGAGAAACCAGCAGAAGC
SEQ2602 TGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATTCCTTTGAGAAACCAGCAGAAGC
SEQ2603 TGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATTCCTTTGAGAAACCAGCAGAAGC
SEQ2604 TGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATTCCTTTGAGAAACCAGCAGAAGC
SEQ2605 TGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATTCCTTTGAGAAACCAGCAGAAGC
SEQ2606 TGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATTCCTTTGAGAAACCAGCAGAAGC
SEQ2607 TGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATTCCTTTGAGAAACCAGCAGAAGC
SEQ2608 TGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATTCCTTTGAGAAACCAGCAGAAGC
SEQ2609 TGTTATTCGTAAAGAGCTCAGTCATTTATCACTAAATTCCTTTGAGAAACCAGCAGAAGC

SEQ2601 ATATAAGGAACGTTTGAAAGAAATACTTGCAAAAGCAAGACAAGATAATCCTAAATTGCC
SEQ2602 ATATAAGGAACGTTTGAAAGAAATCCTTGCAAAAGCAAGACAAGATAATCCTAAATTGCC
SEQ2603 ATATAAGGAACGTTTGAAAGAAATCCTTGCAAAAGCAAGACAAGATAATCCTAAATTGCC
SEQ2604 ATATAAGGAACGTTTGAAAGAAATTCTTGCAAAAGCAAGACAAGATAATCCTAAATTGCC
SEQ2605 ATATAAGGAACGTTTGAAAGAAATACTTGCAAAAGCAAGACAAGATAATCCTAAATTGCC
SEQ2606 ATATAAGGAACGTTTGAAAGAAATTCTTGCAAAAGCAAGACAAGATAATCCTAAATTGCC
SEQ2607 ATATAAGGAACGTTTGAAAGAAATCCTTGCAAAAGCAAGACAAGATAATCCTAAATTGCC
SEQ2608 ATATAAGGAACGTTTGAAAGAAATCCTTGCAAAAGCAAGACAAGATAATCCTAAATTGCC
SEQ2609 ATATAAGGAACGTTTGAAAGAAATTCTTGCAAAAGCAAGACAAGATAATCCTAAATTGCC

SEQ2601 TATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAAT
SEQ2602 TATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAAT
SEQ2603 TATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAAT
SEQ2604 TATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAAT
SEQ2605 TATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAAT
SEQ2606 TATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAAT
SEQ2607 TATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAAT
SEQ2608 TATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAAT
SEQ2609 TATTTATGTTTTAGGCATTTATAATCCTTTTTACCTAAACTTTCCACAATTAACTAAAAT

SEQ2601 GCAAACCGTTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAA
SEQ2602 GCAAACCGTTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAA
SEQ2603 GCAAACCGTTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAA
SEQ2604 GCAAACCGTTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAA
SEQ2605 GCAAACCGTTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAA
SEQ2606 GCAAACCGTTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAA
SEQ2607 GCAAACCGTTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAA
SEQ2608 GCAAACCGTTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAA
SEQ2609 GCAAACCGTTATTGATAATTGGAATAAAGCTACAAAAGAAGTAGTTGATGCTTCAGAAAA

SEQ2601 TGTTTATTTTGTCCCAATTAATGACCGCCTTTATAAGGGAATAAATGGTAAAGAGGGTAT
SEQ2602 TGTTTATTTTGTCCCAATTAATGACCGCCTTTATAAGGGAATAAATGGTAAAGAGGGTAT
SEQ2603 TGTTTATTTTGTCCCAATTAATGACCGCCTTTATAAGGGAATAAATGGTAAAGAGGGTAT
SEQ2604 TGTTTATTTTGTCCCAATTAATGACCGCCTTTATAAGGGAATAAATGGTAAAGAGGGTAT
SEQ2605 TGTTTATTTTGTCCCAATTAATGACCGCCTTTATAAGGGAATAAATGGTAAAGAGGGTAT
SEQ2606 TGTTTATTTTGTCCCAATTAATGACCGCCTTTATAAGGGAATAAATGGTAAAGAGGGTAT
SEQ2607 TGTTTATTTTGTCCCAATTAATGACCGCCTTTATAAGGGAATAAATGGTAAAGAGGGTAT
SEQ2608 TGTTTATTTTGTCCCAATTAATGACCGCCTTTATAAGGGAATAAATGGTAAAGAGGGTAT
SEQ2609 TGTTTATTTTGTCCCAATTAATGACCGCCTTTATAAGGGAATAAATGGTAAAGAGGGTAT

SEQ2601 TACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTACTGGAGACCA
SEQ2602 TATAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTACTGGAGACCA
SEQ2603 TACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTACTGGAGACCA
SEQ2604 TACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTACTGGAGACCA
SEQ2605 TACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTACTGGAGACCA
SEQ2606 TACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTACTGGAGACCA
SEQ2607 TACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTACTGGAGACCA
SEQ2608 TACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTACTGGAGACCA
SEQ2609 TACAGAGTCATCAAATAGTCAGGCAAGTATCACTAATGATGCTCTCTTTACTGGAGACCA

SEQ2601 TTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAA
SEQ2602 TTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAA

Table 26: Comparative Sequences relating to SAG0503 (lipase/acylhydolase)

```
SEQ2603          TTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAA

SEQ2604          TTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAA
SEQ2605          TTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAA
SEQ2606          TTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAA
SEQ2607          TTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAA
SEQ2608          TTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAA
SEQ2609          TTTTCATCCCAATAATATTGGCTATCAAATCATGTCTAACGCCGTTATGGAGAAAATAAA

SEQ2601          TGAAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAAG----------------
SEQ2602          TGAAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAAGTGGTCC----------
SEQ2603          TGAAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAA------------------
SEQ2604          TGAAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAA-----------------
SEQ2605          TGAAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAA------------------
SEQ2606          TGAAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAA-----------------
SEQ2607          TGAAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAA-----------------
SEQ2608          TGAAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAAGTGG-------------
SEQ2609          TGAAACAAGAAAAAACTGGCCGAACCCAGCTTTCTTGTACAAATABCMARATVSTNCSRA

SEQ2601          ----------------
SEQ2602          ----------------
SEQ2603          ----------------
SEQ2604          ----------------
SEQ2605          ----------------
SEQ2606          ----------------
SEQ2607          ----------------
SEQ2608          ----------------
SEQ2609          NGTSAGASACYHYDAS
```

>SEQ ID NO 2650:103_090 frame: 2

```
IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVP
LLSESLHNRYSYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLA
VIRKELSHLSLNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKM
QTVIDNWNKATKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDH
FHPNNIGYQIMSNAVMEKINETRKNWP
```

>SEQ ID NO 2651:103_H36B frame: 2

```
IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLS
ESLHNRYSYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIR
KELSHLSLNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTV
IDNWNKATKEVVDASENVYFVPINDRLYKGINGKEGIIESSNSQASITNDALFTGDHFHP
NNIGYQIMSNAVMEKINETRKNWP
```

>SEQ ID NO 2652:103_18RS21 frame: 3

```
IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLS
ESLHNRYSYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIR
KELSHLSLNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTV
IDNWNKATKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHP
NNIGYQIMSNAVMEKINETRKNWP
```

>SEQ ID NO 2653:103_COH1 frame: 3

```
IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPL
LSESLHNRYSYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAV
IRKELSHLSLNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQ
TVIDNWNKATKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHF
HPNNIGYQIMSNAVMEKINETRKNWP
```

>SEQ ID NO 2654:103_CJB110 frame: 3

```
IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLS
ESLHNRYSYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIR
KELSHLSLNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTV
IDNWNKATKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHP
NNIGYQIMSNAVMEKINETRKNWP
```

>SEQ ID NO 2655:103_1169NT frame: 3

```
IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLS
ESLHNRYSYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIR
KELSHLSLNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTV
IDNWNKATKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHP
NNIGYQIMSNAVMEKINETRKNWP
```

Table 26: Comparative Sequences relating to SAG0503 (lipase/acylhydolase)

```
>SEQ ID NO 2656:103_JM9130013 frame: 3
IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLS
ESLHNRYSYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIR
KELSHLSLNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTV
IDNWNKATKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHP
NNIGYQIMSNAVMEKINETRKNWP
>SEQ ID NO 2657:103_2603 frame: 1
IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLL
SESLHNRYSYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVI
RKELSHLSLNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQT
VIDNWNKATKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFH
PNNIGYQIMSNAVMEKINETRKNWP

>SEQ ID NO 2658:103_M781 frame: 3
IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPL
LSESLHNRYSYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAV
IRKELSHLSLNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQ
TVIDNWNKATKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHF
HPNNIGYQIMSNAVMEKINETRKNWP
```

```
SEQ2650        IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLSESLHNRY
SEQ2651        IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLSESLHNRY
SEQ2652        IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLSESLHNRY
SEQ2653        IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLSESLHNRY
SEQ2654        IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLSESLHNRY
SEQ2655        IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLSESLHNRY
SEQ2656        IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLSESLHNRY
SEQ2657        IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLSESLHNRY
SEQ2658        IFSLIIPKSNPKLTKKDFLTKKVIPLNYVALGDSLTEGVGDTTSQGGFVPLLSESLHNRY

SEQ2650        SYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIRKELSHLS
SEQ2651        SYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIRKELSHLS
SEQ2652        SYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIRKELSHLS
SEQ2653        SYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIRKELSHLS
SEQ2654        SYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIRKELSHLS
SEQ2655        SYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIRKELSHLS
SEQ2656        SYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIRKELSHLS
SEQ2657        SYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIRKELSHLS
SEQ2658        SYQVTSVNYGVSGNTSQQILKRMTTDPQIEKDLEKADLLTLTVGGNDVLAVIRKELSHLS

SEQ2650        LNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTVIDNWNKA
SEQ2651        LNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTVIDNWNKA
SEQ2652        LNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTVIDNWNKA
SEQ2653        LNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTVIDNWNKA
SEQ2654        LNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTVIDNWNKA
SEQ2655        LNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTVIDNWNKA
SEQ2656        LNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTVIDNWNKA
SEQ2657        LNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTVIDNWNKA
SEQ2658        LNSFEKPAEAYKERLKEILAKARQDNPKLPIYVLGIYNPFYLNFPQLTKMQTVIDNWNKA

SEQ2650        TKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHPNNIGYQI
SEQ2651        TKEVVDASENVYFVPINDRLYKGINGKEGIIESSNSQASITNDALFTGDHFHPNNIGYQI
SEQ2652        TKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHPNNIGYQI
SEQ2653        TKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHPNNIGYQI
SEQ2654        TKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHPNNIGYQI
SEQ2655        TKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHPNNIGYQI
SEQ2656        TKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHPNNIGYQI
SEQ2657        TKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHPNNIGYQI
SEQ2658        TKEVVDASENVYFVPINDRLYKGINGKEGITESSNSQASITNDALFTGDHFHPNNIGYQI

SEQ2650        MSNAVMEKINETRKNWP
SEQ2651        MSNAVMEKINETRKNWP
SEQ2652        MSNAVMEKINETRKNWP
SEQ2653        MSNAVMEKINETRKNWP
SEQ2654        MSNAVMEKINETRKNWP
SEQ2655        MSNAVMEKINETRKNWP
SEQ2656        MSNAVMEKINETRKNWP
SEQ2657        MSNAVMEKINETRKNWP
SEQ2658        MSNAVMEKINETRKNWP
```

Table 27: Comparative Sequences relating to SAG1473 (cell wall surface anchor family protein)

SEQ ID NO. 2701: SAG1473 FROM THE 1169NT1 GBS TYPE V STRAIN (REVERSE COMPLEMENT)
GATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATCAGATGA
ACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCATCAAGTGAACCAGAAACAAATC
CGTCAACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGTAGA
ACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTAT
TAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAA
GTGATGGGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

SEQ ID NO. 2702: SAG1473 FROM THE 18RS21 GBS TYPE II STRAIN
GATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATCAGATGA
ACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCATCAAGTGAACCAGAAACAAATC
CGTCAACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGTAGA
ACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTAT
TAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAA
ATGATGGGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

SEQ ID NO. 2703: SAG1473 FROM THE 2603 V/R GBS TYPE V STRAIN
GATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATCAGATGA
ACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCATCAAGTGAACCAGAAACAAATC
CGTCAACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGTAGA
ACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTAT
TAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAA
ATGATGGGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

SEQ ID NO. 2704: SAG1473 FROM THE 090 GBS TYPE Ia STRAIN
GACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCATCAAGTGAACCAGAAACAAATCCGTC
AACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGTAGAACGA
AGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAG
AATTTTAGTAAAGCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAAATGA
TGGGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

SEQ ID NO. 2705: SAG1473 FROM THE A909 GBS TYPE Ia STRAIN
GATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATTAGATGA
ACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCATCAAGTGAACCAGAAACAAATC
CCTCAACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGTAGC
ACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTAT
TAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAA
ATGATGAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

SEQ ID NO. 2706: SAG1473 FROM THE CJB110 GBS NONTYPEABLE STRAIN
GATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATCAGATGA
ACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCATCAAGTGAACCAGAAACAAATC
CGTCAACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGTAGA
ACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTAT
TAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAA
ATGATGGGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

Table 27: Comparative Sequences relating to SAG1473
(cell wall surface anchor family protein)

SEQ ID NO. 2707: SAG1473 FROM THE COH1 GBS TYPE III STRAIN
(REVERSE COMPLEMENT)
GATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATCAGATGA
ACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCAAGTTCATCAAGTGAACCAGAAACAAATC
CCTCAACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGGAGC
ACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTAT
TAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAAGAAGTGGAACGCGATGAATCATCATCTTCAAAAGCAA
ATGATGAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

SEQ ID NO. 2708: SAG1473 FROM THE H36b GBS TYPE Ib STRAIN
GATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATTAGATGA
ACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCATCAAGTGAACCAGAAACAAATC
CCTCAACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGTAGC
ACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTAT
TAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAArAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAA
ATGATGAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

SEQ ID NO. 2709: SAG1473 FROM THE JM910013 GBS TYPE VIII STRAIN
GATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATTAGATGA
ACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCATCAAGTGAACCAGAAACAAATC
CCTCAACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGTAGC
ACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTAT
TAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAA
ATGATGAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

SEQ ID NO. 2710: SAG1473 FROM THE M732 GBS TYPE III STRAIN
GATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATCAGATGA
ACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCAAGTTCATCAAGTGAACCAGAAACAAATC
CCTCAACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGGAGC
ACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTAT
TAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAAGAAGTGGAACGCGATGAATCATCATCTTCAAAAGCAA
ATGATGAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

SEQ ID NO. 2711: SAG1473 FROM THE M781 GBS TYPE III STRAIN
GATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATCAGATGA
ACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCAAGTTCATCAAGTGAACCAGAAACAAATC
CCTCAACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGATGGGAGC
ACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTAT
TAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAA
ATGATGAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA

```
SEQ2701          ATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAA
SEQ2702          ATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAA
SEQ2703          ATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAA
SEQ2704          -----------------------------------------------------------
SEQ2705          ATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAA
SEQ2706          ATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAA
SEQ2707          ATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAA
SEQ2709          ATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAA
SEQ2710          ATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAA
SEQ2711          ATACAAGTGATAAGAATACTGACACGAGTGTCGTGACTACGACCTTATCTGAGGAGAAA
```

Table 27: Comparative Sequences relating to SAG1473
(cell wall surface anchor family protein)

SEQ2701	GATCAGATGAACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCA
SEQ2702	GATCAGATGAACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCA
SEQ2703	GATCAGATGAACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCA
SEQ2704	--------------GACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCA
SEQ2705	GATTAGATGAACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCA
SEQ2706	GATCAGATGAACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCA
SEQ2707	GATCAGATGAACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCAAGTTCA
SEQ2709	GATTAGATGAACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCGAGTTCA
SEQ2710	GATCAGATGAACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCAAGTTCA
SEQ2711	GATCAGATGAACTAGACCAGTCTAGTACTGGTTCTTCTTCTGAAAATGAATCAAGTTCA

SEQ2701	TCAAGTGAACCAGAAACAAATCCGTCAACTAATCCACCTACAACAGAACCATCGCAACCC
SEQ2702	TCAAGTGAACCAGAAACAAATCCGTCAACTAATCCACCTACAACAGAACCATCGCAACCC
SEQ2703	TCAAGTGAACCAGAAACAAATCCGTCAACTAATCCACCTACAACAGAACCATCGCAACCC
SEQ2704	TCAAGTGAACCAGAAACAAATCCGTCAACTAATCCACCTACAACAGAACCATCGCAACCC
SEQ2705	TCAAGTGAACCAGAAACAAATCCCTCAACTAATCCACCTACAACAGAACCATCGCAACCC
SEQ2706	TCAAGTGAACCAGAAACAAATCCGTCAACTAATCCACCTACAACAGAACCATCGCAACCC
SEQ2707	TCAAGTGAACCAGAAACAAATCCCTCAACTAATCCACCTACAACAGAACCATCGCAACCC
SEQ2709	TCAAGTGAACCAGAAACAAATCCCTCAACTAATCCACCTACAACAGAACCATCGCAACCC
SEQ2710	TCAAGTGAACCAGAAACAAATCCCTCAACTAATCCACCTACAACAGAACCATCGCAACCC
SEQ2711	TCAAGTGAACCAGAAACAAATCCCTCAACTAATCCACCTACAACAGAACCATCGCAACCC

SEQ2701	TCACCTAGTGAAGAGAACAAGCCTGATGGTAGAACGAAGACAGAAATTGGCAATAATAAG
SEQ2702	TCACCTAGTGAAGAGAACAAGCCTGATGGTAGAACGAAGACAGAAATTGGCAATAATAAG
SEQ2703	TCACCTAGTGAAGAGAACAAGCCTGATGGTAGAACGAAGACAGAAATTGGCAATAATAAG
SEQ2704	TCACCTAGTGAAGAGAACAAGCCTGATGGTAGAACGAAGACAGAAATTGGCAATAATAAG
SEQ2705	TCACCTAGTGAAGAGAACAAGCCTGATGGTAGCACGAAGACAGAAATTGGCAATAATAAG
SEQ2706	TCACCTAGTGAAGAGAACAAGCCTGATGGTAGAACGAAGACAGAAATTGGCAATAATAAG
SEQ2707	TCACCTAGTGAAGAGAACAAGCCTGATGGGAGCACGAAGACAGAAATTGGCAATAATAAG
SEQ2709	TCACCTAGTGAAGAGAACAAGCCTGATGGTAGCACGAAGACAGAAATTGGCAATAATAAG
SEQ2710	TCACCTAGTGAAGAGAACAAGCCTGATGGGAGCACGAAGACAGAAATTGGCAATAATAAG
SEQ2711	TCACCTAGTGAAGAGAACAAGCCTGATGGGAGCACGAAGACAGAAATTGGCAATAATAAG

SEQ2701	GATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAGAATTTTAGTAAA
SEQ2702	GATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAGAATTTTAGTAAA
SEQ2703	GATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAGAATTTTAGTAAA
SEQ2704	GATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAGAATTTTAGTAAA
SEQ2705	GATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAGAATTTTAGTAAA
SEQ2706	GATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAGAATTTTAGTAAA
SEQ2707	GATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAGAATTTTAGTAAA
SEQ2709	GATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAGAATTTTAGTAAA
SEQ2710	GATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAGAATTTTAGTAAA
SEQ2711	GATATTTCTAGTGGAACAAAAGTATTAATTTCAGAAGATAGTATTAAGAATTTTAGTAAA

SEQ2701	GCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAAGTGAT
SEQ2702	GCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAAATGAT
SEQ2703	GCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAAATGAT
SEQ2704	GCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAAATGAT
SEQ2705	GCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAAATGAT
SEQ2706	GCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAAATGAT
SEQ2707	GCAAGTAGTGATCAAGAAGAAGTGGAACGCGATGAATCATCATCTTCAAAAGCAAATGAT
SEQ2709	GCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAAATGAT
SEQ2710	GCAAGTAGTGATCAAGAAGAAGTGGAACGCGATGAATCATCATCTTCAAAAGCAAATGAT
SEQ2711	GCAAGTAGTGATCAAGAAGAAGTGGATCGCGATGAATCATCATCTTCAAAAGCAAATGAT

Table 27: Comparative Sequences relating to SAG1473 (cell wall surface anchor family protein)

```
SEQ2701          GGGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA---------------------------
SEQ2702          GGGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA---------------------------
SEQ2703          GGGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA---------------------------
SEQ2704          GGGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA---------------------------
SEQ2705          GAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA---------------------------
SEQ2706          GGGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA---------------------------
SEQ2707          GAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAASGATACAAGTGATAAGAATACTGACAC
SEQ2709          GAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA---------------------------
SEQ2710          GAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAA---------------------------
SEQ2711          GAGAAAAAAGGCCACAGTAAGCCTAAAAAGGAATABCMARATVSTNCSRATNGTSAGCWA

SEQ2701          ------------------------------------------------------------
SEQ2702          ------------------------------------------------------------
SEQ2703          ------------------------------------------------------------
SEQ2704          ------------------------------------------------------------
SEQ2705          ------------------------------------------------------------
SEQ2706          ------------------------------------------------------------
SEQ2707          AGTGTCGTGACTACGACCTTATCTGAGGAGAAAAGATTAGATGAACTAGACCAGTCTAG
SEQ2709          ------------------------------------------------------------
SEQ2710          ------------------------------------------------------------
SEQ2711          TRACANCHRAMYRTN---------------------------------------------

SEQ2701          ------------------------------------------------------------
SEQ2702          ------------------------------------------------------------
SEQ2703          ------------------------------------------------------------
SEQ2704          ------------------------------------------------------------
SEQ2705          ------------------------------------------------------------
SEQ2706          ------------------------------------------------------------
SEQ2707          ACTGGTTCTTCTTCTGAAAATGAATCGAGTTCATCAAGTGAACCAGAAACAAATCCCTC
SEQ2709          ------------------------------------------------------------
SEQ2710          ------------------------------------------------------------
SEQ2711          ------------------------------------------------------------

SEQ2701          ------------------------------------------------------------
SEQ2702          ------------------------------------------------------------
SEQ2703          ------------------------------------------------------------
SEQ2704          ------------------------------------------------------------
SEQ2705          ------------------------------------------------------------
SEQ2706          ------------------------------------------------------------
SEQ2707          ACTAATCCACCTACAACAGAACCATCGCAACCCTCACCTAGTGAAGAGAACAAGCCTGA
SEQ2709          ------------------------------------------------------------
SEQ2710          ------------------------------------------------------------
SEQ2711          ------------------------------------------------------------

SEQ2701          ------------------------------------------------------------
SEQ2702          ------------------------------------------------------------
SEQ2703          ------------------------------------------------------------
SEQ2704          ------------------------------------------------------------
SEQ2705          ------------------------------------------------------------
SEQ2706          ------------------------------------------------------------
SEQ2707          GGTAGCACGAAGACAGAAATTGGCAATAATAAGGATATTTCTAGTGGAACAAAAGTATT
SEQ2709          ------------------------------------------------------------
SEQ2710          ------------------------------------------------------------
SEQ2711          ------------------------------------------------------------
```

Table 27: Comparative Sequences relating to SAG1473 (cell wall surface anchor family protein)

```
SEQ2701      ------------------------------------------------------------
SEQ2702      ------------------------------------------------------------
SEQ2703      ------------------------------------------------------------
SEQ2704      ------------------------------------------------------------
SEQ2705      ------------------------------------------------------------
SEQ2706      ------------------------------------------------------------
SEQ2707      ATTTCAGAAGATAGTATTAAGAATTTTAGTAAAGCAAGTAGTGATCAAGAARAAGTGGA
SEQ2709      ------------------------------------------------------------
SEQ2710      ------------------------------------------------------------
SEQ2711      ------------------------------------------------------------

SEQ2701      ------------------------------------------------------------
SEQ2702      ------------------------------------------------------------
SEQ2703      ------------------------------------------------------------
SEQ2704      ------------------------------------------------------------
SEQ2705      ------------------------------------------------------------
SEQ2706      ------------------------------------------------------------
SEQ2707      CGCGATGAATCATCATCTTCAAAAGCAAATGATGAGAAAAAAGGCCACAGTAAGCCTAA
SEQ2709      ------------------------------------------------------------
SEQ2710      ------------------------------------------------------------
SEQ2711      ------------------------------------------------------------

SEQ2701      ------
SEQ2702      ------
SEQ2703      ------
SEQ2704      ------
SEQ2705      ------
SEQ2706      ------
SEQ2707      AAGGAA
SEQ2709      ------
SEQ2710      ------
SEQ2711      ------
```

>SEQ ID NO 2750:4_1169NT frame: 1
DTSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SPSEENKPDGRTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKASD
GKKGHSKPKKE

>SEQ ID NO 2751:4_18RS21 frame: 1
DTSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SPSEENKPDGRTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
GKKGHSKPKKE

>SEQ ID NO 2752:4_2603 frame: 1
DTSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SPSEENKPDGRTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
GKKGHSKPKKE

>SEQ ID NO 2753:4_090 frame: 1
DQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQPSPSEENKPDGRTKTEIGNNKDISSG
TKVLISEDSIKNFSKASSDQEEVDRDESSSSKANDGKKGHSKPKKE

>SEQ ID NO 2754:4_A909 frame: 1
DTSDKNTDTSVVTTTLSEEKRLDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
EKKGHSKPKKE

>SEQ ID NO 2755:4_CJB110 frame: 1
DTSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SPSEENKPDGRTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
GKKGHSKPKKE

Table 27: Comparative Sequences relating to SAG1473 (cell wall surface anchor family protein)

>SEQ ID NO 2756:4_COH1 frame: 1
DTSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVERDESSSSKAND
EKKGHSKPKKE

>SEQ ID NO 2757:4_H36B frame: 1
DTSDKNTDTSVVTTTLSEEKRLDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEXVDRDESSSSKAND
EKKGHSKPKKE

>SEQ ID NO 2758:4_JM9130013 frame: 1
DTSDKNTDTSVVTTTLSEEKRLDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
EKKGHSKPKKE

>SEQ ID NO 2759:4_M732 frame: 1
DTSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVERDESSSSKAND
EKKGHSKPKKE

>SEQ ID NO 2760:4_M781 frame: 1
DTSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
EKKGHSKPKKE

```
SEQ2750         TSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SEQ2751         TSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SEQ2752         TSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SEQ2753         ------------------------DQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SEQ2754         TSDKNTDTSVVTTTLSEEKRLDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SEQ2755         TSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SEQ2756         TSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SEQ2757         TSDKNTDTSVVTTTLSEEKRLDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SEQ2758         TSDKNTDTSVVTTTLSEEKRLDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SEQ2759         TSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP
SEQ2760         TSDKNTDTSVVTTTLSEEKRSDELDQSSTGSSSENESSSSSEPETNPSTNPPTTEPSQP

SEQ2750         SPSEENKPDGRTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKASD
SEQ2751         SPSEENKPDGRTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
SEQ2752         SPSEENKPDGRTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
SEQ2753         SPSEENKPDGRTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
SEQ2754         SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
SEQ2755         SPSEENKPDGRTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
SEQ2756         SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVERDESSSSKAND
SEQ2757         SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEXVDRDESSSSKAND
SEQ2758         SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND
SEQ2759         SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVERDESSSSKAND
SEQ2760         SPSEENKPDGSTKTEIGNNKDISSGTKVLISEDSIKNFSKASSDQEEVDRDESSSSKAND

SEQ2750         KKGHSKPKKE
SEQ2751         KKGHSKPKKE
SEQ2752         KKGHSKPKKE
SEQ2753         KKGHSKPKKE
SEQ2754         KKGHSKPKKE
SEQ2755         KKGHSKPKKE
SEQ2756         KKGHSKPKKE
SEQ2757         KKGHSKPKKE
SEQ2758         KKGHSKPKKE
SEQ2759         KKGHSKPKKE
SEQ2760         KKGHSKPKKE
```

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

SEQ ID NO. 2801: SAG1552 FROM THE 1169NT1 GBS TYPE V STRAIN (REVERSE COMPLEMENT)

TTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTTAAAGGAGTAGACGTTGAGTCT
TCCTTAGCAGGTTATCATCACAACGATTTTCCTATTACTCAAAAAACGTATCGTGAGTGGTTCCATTTAATTTCCAAC
ATGGGGGCAAATACTGTAAGAGTCAAAGTACCGATGAATGTTGCATTTTACGATGCTTTATATCACCACAACAAAGCA
TCAAAGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCTATAACAGCTTTTAAT
GATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTGGATATTCTCCATGGGCGTAAGCAAGTATGGAAT
ACTGATTTTGGTAGCCGTCATTATCATTATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAAT
AGTGGTACTGTCGCTTATACTAATCATCAAGAGAAAAAAACGCAATATAAAGGACGTTATTTTAAAACTTCTGCGGCA
GCTAATCCATTTGAGGTCATGCTAGCTCAAGTTATGGATGAATTGACACATTATGAGACAGCTAAATATGGTTGGCAA
CATTTGATTAGTTTTTCAAACTCACCAACAACAGACCCTTTTCGTTATCGAAAACCATTTGAGGCACAGGCTCCTAAA
TACGTACAACTAAATGTAGAAAATATTCAAGCTAATTCGAATGTTAAAGCAGGTATTTTTGCAGCATATAAAGCTATT
GATTTCCATCCTCGATACAAGGATTATCTATTATTTGATAAAGAGAATATCAGTAAAGAAGATAGACAAAAGATTAAA
GAACTTTCTTTGTCACAGGGATACGTTAAACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGC
TATTCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGAAAAAGAACAAGGTCAG
CGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTTTGGAGCGACTATCAATGCATGGCAAGACGATTGG
AATGCAAGGGCGTGGAATACATCCTTCGCCACAAATAAACATAGTCAATTCCTATGGGGGATGCACAAGTATTTAAT
CAAGGTTATGGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAAAGGAGAGTGG
AAACATCCTCTG

SEQ ID NO. 2802: SAG1552 FROM THE __________

ATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAGTGATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCT
GAAAAACTAAAAGAAAAACGATTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAG
GTCACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATTATTTGTCCAAGAGCGC
TATAATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACGGTAAAGATTTTTATGCTTTCCCACCAAAGAAGAACAGT
AGTAATTTTGAGCAGATCAATATGGTATTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAG
AGGTTCTTACCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAATTGATAGGCACCAAAAAACATTTGATTCACAA
ACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCAGTTGTTGAATTTTTCTGATCCATCATCT
CAAAAAATTCACGATGATTACTTTAAACATTATGGTGTGAAGGAGTTAGAAATTGAGAGCATTGCTTTAGGATTAGGT
GCTAATAGCAAAGAAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACCAAAACC
TTTTTAAAAGACTCCTATTATAGTATTTAAGAAAGAA

SEQ ID NO. 2803: SAG1552 FROM THE 18RS21 GBS TYPE II STRAIN

AAGGGCTTATTAAAAGAAAATACAAGAACTAACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAA
CCTTTTGTTGTTAAAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACTCAAAAA
ACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGGCAAATACTGTAAGAGTCAAGGTACCGATGAATGTTGCA
TTTTACGATGCCTTATATCACCACAACAAAGCATCAAAGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCT
TATCGCAATAATGCTTCTATAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
GATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTGGGTAGCCGTCATTATCATTATGATCTTAGTCCTTGG
GTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGTACTGTCGCTTATACTAATCATCAAGAGAAAAAAACGCAA
TATAAAGGACGTTATTTTAAAACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGATGAATTG
ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCAACAACAGACCCTTTTCAT
TATCGAAAACCATTTGAGGCACAGGCTCCTAAATACGTACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTT
AAAGCAGGTATGTTTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATAAAGAG
AATATCAGTAAAGAAGATAGACAAAAGATTAAAGAACTTTCTTTGTCACAGGGATACGTTAAACTGCTAAATGCTTAT
CACAAAATCCCTGTTCTAGTCACGGGTTATGGCTATTCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGT
CCTCTGCCGATTAATGAAAAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTTT
GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTTCGCCACAAATAAACATAGT
CAATTCCTATGGGGGATGCACAAGTATTTAATCAAGGTTATGGTTTATTAGGCTTTAAAAACGCAAAACATCATTAT
CAAGTTGATGGTAAAAGAGGCAAAGGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCT
AGCAGTGATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACGATTATTACCAATA
GATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGTCACATTTTCTAAATCTAGTGACTTTGTATTG
TCTATTGATCCAAATGGCAAGTCTGAATTATTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAG
CTTAACGGTAAAGATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGTATTGAGA
AATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTTACCAACTCATCCTACTGGTCTTCTC
AAAACAGGAACAACTGATAGGCACCAAAAAACATTTGATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAG
GTCAGAATTCCGTGGCAGTTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTAT
GGTGTGAAGGAGTTAGAAATTGAGAGCATTGCTTTAGGATTAGGTGCTAATAGCAAAGAAAACACACTGATAAAGATG
GCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACCAAAACCTTTTTAAAAGACTCCTATTATGTATTAAGAAAG
AA

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

SEQ ID NO. 2804: SAG1552 FROM THE 2603 V/R GBS TYPE V STRAIN
(REVERSE COMPLEMENT)

TATTAAAAGAAAATACAAGAACTAACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTG
TTGTTAAAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACTCAAAAAACGTATC
GTGAATGGTTCCATTTAATTTCCAACATGGGGGCAAATACTGTAAGAGTCAAGGTACCGATGAATGTTGCATTTTACG
ATGCCTTATATCACCACAACAAAGCATCAAAGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCA
ATAATGCTTCTATAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTGGATATTC
TCCATGGGCGTAAGCAAGTATGGAATACTGATTTGGGTAGCCGTCATTATCATTATGATCTTAGTCCTTGGGTACTTG
GTTATGTCGTAGGGGATGATTGGAATAGTGGTACTGTCGCTTATACTAATCATCAAGAGAAAAAAACGCAATATAAAG
GACGTTATTTTAAAACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGATGAATTGACACATT
ATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCAACAACAGACCCTTTTCATTATCGAA
AACCATTTGAGGCACAGGCTCCTAAATACGTACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTTAAAGCAG
GTATGTTTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATAAAGAGAATATCA
GTAAAGAAGATAGACAAAAGATTAAAGAACTTTCTTTGTCACAGGGATACGTTAAACTGCTAAATGCTTATCACAAAA
TCCCTGTTCTAGTCACGGGTTATGGCTATTCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGC
CGATTAATGAAAAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTTTGGAGCGA
CTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTTCGCCACAAATAAACATAGTCAATTCC
TATGGGGGGATGCACAAGTATTTAATCAAGGTTATGGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTG
ATGGTAAAAGAGGCAAAGGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAGTG
ATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACGATTATTACCAATAGATATTA
CACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGTCACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTG
ATCCAAATGGCAAGTCTGAATTATTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACG
GTAAAGATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGTATTGAGAAATACAA
AGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTTACCAACTCATCCTACTGGTCTTCTCAAAACAG
GAACAACTGATAGGCACCAAAAAACATTTGATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAA
TTCCGTGGCAGTTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTATGGTGTGA
AGGAGTTAGAAATTGAGAGCATTGCTTTAGGATTAGGTGCTAATAGCAAAGAAAACACACTGATAAAGATGGCAGATT
ATCGTTTGAAAAATTGGGAGAGACCCGATACCAAAACCTTTTTAAAAGACTCCTATTATAGTATTAAGAAAGAATGGT
CTAAAGAAAGAGAGAGAACATATGGTCCA

SEQ ID NO. 2805: SAG1552 FROM THE A909 GBS TYPE Ia STRAIN
(REVERSE COMPLEMENT)

AAGGGCTTATTAAAAGAAAATACAAGAACTAACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAA
CCTTTTGTTGTTAAAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACTCAAAAA
ACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGGCAAATACTGTAAGAGTCAAGGTACCGATGAATGTTGCA
TTTTACGATGCCTTATATCACCACAACAAAGCATCAAAGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCT
TATCGCAATAATGCTTCTATAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
GATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTGGGTAGCCGTCATTATCATTATGATCTTAGTCCTTGG
GTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGTACTGTCGCTTATACTAATCATCAAGAGAAAAAAACGCAA
TATAAAGGACGTTATTTTAAAACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGATGAATTG
ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCAACAACAGACCCTTTTCAT
TATCGAAAACCATTTGAGGCACAGGCTCCTAAATACGTACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTT
AAAGCAGGTATGTTTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATAAAGAG
AATATCAGTAAAGAAGATAGACAAAAGATTAAAGAACTTTCTTTGTCACAGGGATACGTTAAACTGCTAAATGCTTAT
CACAAAATCCCTGTTCTAGTCACGGGTTATGGCTATTCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGT
CCTCTGCCGATTAATGAAAAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTTT
GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTTCGCCACAAATAAACATAGT
CAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTATGGTTTATTAGGCTTTAAAAACGCAAAACATCATTAT
CAAGTTGATGGTAAAAGAGGCAAAGGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCT
AGCAGTGATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACGATTATTACCAATA
GATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGTCACATTTTCTAAATCTAGTGACTTTGTATTG
TCTATTGATCCAAATGGCAAGTCTGAATTATTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAG
CTTAACGGTAAAGATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGTATTGAGA
AATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTTACCAACTCATCCTACTGGTCTTCTC
AAAACAGGAACAACTGATAGGCACCAAAAAACATTTGATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAG
GTCAGAATTCCGTGGCAGTTGTTGAATTTTTCTGATCCATCATCTCAAAGAATTCACGATGATTACTTTAAACATTAT
GGTGTGAAGGAGTTAGAAAATTGAGAGCCATTGCTTTAGGATTAGGTGCTAATAGCAAAGAAAACACACTGATAAAGA
TGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACCAAAACCTTTTTAAAAGA

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

SEQ ID NO. 2806: SAG1552 FROM THE CJB110 GBS NONTYPEABLE STRAIN

TATTACTTTGATGGTAGTTTGTATTTACCAAAGGGCTTATTAAAAGAAAATACAAGAACTAACTTTGTTGTTAAAGGT
GATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTTAAAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTAT
CATCACAACGATTTTCCTATTACTCAAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGGCAAATACT
GTAAGAGTCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGCATCAAAGAGGCCACTG
TATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCTATAACAGCTTTTAATGATAATTATAGGGGG
TATTTAAAACGAGAAGCAAAAGGCGTTGTGGATATTCTCCATGGGCGTAAGCAAGTATGGAATACAGATTTTGGTAGC
CGTCATTATCATTATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGTACTGTCGCT
TATACTAATCATCAAGAGAAAAAAACGCAATATAAAGGACGTTATTTTAAAACTTCTGTGGCAGCTAATCCATTTGAG
GTCATGCTAGCTCAAGTAATGGATGAATTGACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTT
TCAAACTCACCAACAACAGACCCTTTTCATTATCGAAAACCATTTGAGGCACAGGCTCCTAAATACGTACAACTAAAT
GTAGAAAATATTCAAGCTAATTCAAATGTTAAAGCAGGTATGTTTGCAGCATATAAAGCTATTGATTTCCATCCTCGA
TACAAGGATTATCTATTATTTGATAAAGAGAATATCAGTAAAGAAGATAGACAAAAGATTAAAGAACTTTCTTTGTCA
CAGGGATACGTTAAACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTATTCGACAGCGAGA
GGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGAAAAAGAACAAGGTCAGCGTTTACTAGAAGAT
TATGAATCTTTTATATCATCCGGTAGTTTTGGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGG
AATACATCTTTCGCCACAAATAAACATAATCAATTCCTATGGGGGATGCACAAGTATTTAATCAAGGTTATGGTTTA
TTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAAAGGAGAGTGGAAACATCCTCTGATG
ACTAGTGCAACAGGAGATGACTTATATGCTAGCAGTGATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAA
AAACTAAAAGAAAAACGATTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGTC
ACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATTATTTGTCCAAGAGCGCTAT
AATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACGGTAAAGATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGT
AATTTTGAGCAGATAAATATGGTATTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGG
TTCTTACCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATTTGATTCACAAACA
GATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCAGTTGTTGAATTTTTCTGATCCATCATCTCAA
AAAATTCACGATGATTACTTTAAACATTATGGTGTGAAGGAGTTAGAAATTGAGAGCATTGCTTTAGGATTAGGTGCT
AATAGCAAAGAAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACCAAAACCTTT
TTAAAAGACTCCTATTATGTATTAAGAAAGA

SEQ ID NO. 2807: SAG1552 FROM THE COH1 GBS TYPE III STRAIN

TTTACCACAGGGCTTATTAAAAGAAAATACAAGAACTAACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCAC
CAATAAACCTTTTGTTGTTAAAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTAC
TCAAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGGCAAATACTGTAAGAGTCAAGGTACCGATGAA
TGTTGCATTTTACGATGCCTTATATCACCACAACAAAGAATCAAAGAGGCCACTGTATTTGTTGCAAGGAATACGTAT
AGATTCTTATCGCAATAATGCTTCTATAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGG
CGTTGTGGATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTTGGTAGCCGTCATTATCATTATGATCTTAG
TCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGTACTGTCGCTTATACTAATCATCAAGAGAAAAA
AACGCAATATAAAGGACGTTATTTTAAAACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGA
TGAATTGACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCAACAACAGACCC
TTTTCATTATCGAAAACCATTTGAGGCACAGGCTCCTAAATACGTACAACTAAATGTAGAAAATATTCAAGCTAATTC
AAATGTTAAAGCAGGTATGTTTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGA
TAAAGAGAATATCAGTAAAGAAGATAGACAAAAGATTAAAGAACTTTCTTTGTCACAGGGATACGTTAAACTGCTAAA
TGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTATTCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGA
TAAACGTCCTCTGCCGATTAATGAAAAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGG
TAGTTTTGGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTTCGCCACAAATAA
ACATAGTCAATTCCTATGGGGGATGCACAAGTATTTAATCAAGGTTATGGTTTATTAGGCTTTAAAAACGCAAAACA
TCATTATCAAGTTGATGGTAAAAGAGGCAAAGGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTT
ATATGCTAGCAGTGATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACGATTATT
ACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGTCACATTTTCTAAATCTAGTGACTT
TGTATTGTCTATTGATCCAAATGGCAAGTCTGAATTATTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCT
TCGACAGCTTAACGGTAAAGATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGT
ATTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTTACCAACTCATCCTACTGG
TCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATTTGATTCACAACCAGATATTTCGTTTGGAAAGGACTT
TATAGAGGTCAGAATTCCGTGGCAGTTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAA
ACATTATGGTGTGAAGGAGTTAGAAATTGAGAGCATTGCTTTAGGATTAGGTGCTAATAGCAAAGAAAACACACTGAT
AAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACCAAAACCTTTTTAAAAGACT

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

SEQ ID NO. 2808: SAG1552 FROM THE H36b GBS TYPE Ib STRAIN
AAGGGGCTTATTAAAAGAAAATACAAGAACTAACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAA
ACCTTTTGTTGTTAAAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACTCAAAA
AACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGGCAAATACTGTAAGAGTCAAGGTACCGATGAATGTTGC
ATTTTACGATGCCTTATATCACCACAACAAAGCATCAAAGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTC
TTATCGCAATAATGCTTCTATAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGT
GGATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTTGGTAGCAGTCATTATCATTATGATCTTAGTCCTTG
GGTACTTGGTTATGTCGTAGGGGATGATGGACATAGTGGTACTGTCGCTTTATACTAATCATCAAGAGGAGAAAAACG
CAATATAAAGGACGTTATTTTAAAACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGATGAA
TTGACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCAACAACAGACCCTTTT
CATTATCGAAAACCATTTGAGGCACAGGCTCCTAAATACGTACAACTAAATGTAGAAAATATTCAAGCTAATTCGAAT
GTTAAAGCAGGTATGTTTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATAAA
GAGAATATCAGTAAAGAAGATAGACAAAAGATTAAAGAACTTTCTTTGTCACAGGGATACGTTAAACTGCTAAATGCT
TATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTACTCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAA
CGTCCTCTGCCGATTAATGAAAAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGT
TTTGGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGTGTGGAATACATCCTTCGCCACAAATAAACAT
AGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTATGGTTTATTAGGCTTTAAAAACGCAAAACATCAT
TATCAGGTTGATGGTAAAAGAGGCAAAGAAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATAT
GCTAGCAGTGATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACGATTATTACCA
ATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGTCACATTTTCTAAATCTAGTGACTTTGTA
TTGTCTATTGATCCAAATGGCAAGTCTGAATTATTTGTCCAAGAGCGCTATAACGCCTTAAAAGCGAACTATCTTCGA
CAGCTTAATGGTAAAGATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGTATTG
AGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTTACCAACTCATCCTACTGGTCTT
CTCAAAACAGGAACAACTGATAGGCACCAAAAAACATTTGATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATA
GAGGTCAGAATTCCGTGGCAGTTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACAT
TATGGTGTGAAGGAGTTAGAAATTGAGAGCATTGCTTTAGGATTAGGTGCTAATAGCAAAGAAAACACACTGATAAAG
ATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACCAAAACCTTTTTAAAAGACTCCTATTATAGT

SEQ ID NO. 2809: SAG1552 FROM THE JM9130013 GBS TYPE VIII STRAIN
ACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTTAAAGGAGTAGACGTTGAGT
CTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACTCAAAAAACGTATCGTGAATGGTTCCATTTAATTTCCA
ACATGGGGGCAAATACTGTAAGAGTCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAG
CATCAAAGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCTATAACAGCTTTTA
ATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTGGATATTCTCCATGGGCGTAAGCAAGTATGGA
ATACTGATTTTGGTAGCAGTCATTATCATTATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGA
ATAGTGGTACTGTCGCTTATACTAATCATCAAGAGAAAAAAACGCAATATAAAGGACGTTATTTTAAAACTTCTGTGG
CAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGATGAATTGACACATTATGAGACAGCTAAATATGGTTGGC
AACATTTGATTAGTTTTTCAAACTCACCAACAACAGACCCTTTTCATTATCGAAAACCATTTGAGGCACAGGCTCCTA
AATACGTACAACTAAATGTAGAAAATATTCAAGCTAATTCGAATGTTAAAGCAGGTATGTTTGCAGCATATAAAGCTA
TTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATAAAGAGAATATCAGTAAAGAAGATAGACAAAAGATTA
AAGAACTTTCTTTGTCACAGGGATACGTTAAACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATG
GCTACTCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGAAAAAGAACAAGGTC
AGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTTTGGAGCGACTATCAATGCATGGCAAGACGATT
GGAATGCAAGGGTGTGGAATACATCCTTCGCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTA
ATCAAGGTTATGGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAGGTTGATGGTAAAAGAGGCAAAGAAGAGT
GGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAGTGATGAAAGCTATCTCTACCTTGCGA
TTAAAACAAAACCTGAAAAACTAAAAGAAAAACGATTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAA
TGAATGGTAGTAAGGTCACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATTAT
TTGTCCAAGAGCGCTATAACGCCTTAAAAGCGAACTATCTTCGACAGCTTAATGGTAAAGATTTTTATGCTTTCCCAC
CAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGTATTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAG
TAAAAGCAACAGAGAGGTTCTTACCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAA
CATTTGATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCAGTTGTTGAATTTTT
CTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTATGGTGTGAAGGAGTTAGAAATTGAGAGCATTG
CTTTAGGATTAGGTGCTAATAGCAAAGAAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGAC
CCGATACCAAAACCTTTTTAAAAGACTCCTATTATAGTATTAAGAAAG

Table 28: Comparative Sequences relating to SAG1552
(conserved hypothetical protein)

SEQ ID NO. 2810: SAG1552 FROM THE M732 GBS TYPE III STRAIN
TACAAGAACTAACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTTAAAGGAGT
AGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACTCAAAAAACGTATCGTGAATGGTTCCA
TTTAATTTCCAACATGGGGGCAAATACTGTAAGAGTCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCA
CCACAACAAAGAATCAAAGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCTAT
AACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTGGATATTCTCCATGGGCGTAA
GCAAGTATGGAATACTGATTTTGGTAGCCGTCATTATCATTATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGG
GGATGATTGCAATAGTGGTACTGTCGCTTATACTAATCATCAAGAGAAAAAAACGCAATATAAAGGACGTTATTTTAA
AACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGATGAATTGACACATTATGAGACAGCTAA
ATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCAACAACAGACCCTTTTCATTATCGAAAACCATTTGAGGC
ACAGGCTCCTAAATACGTACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTTAAAGCAGGTATGTTTGCAGC
ATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATAAAGAGAATATCAGTAAAGAAGATAG
ACAAAAGATTAAAGAACTTTCTTTGTCACAGGGATACGTTAAACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGT
CACGGGTTATGGCTATTCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGAAAA
AGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTTTGGAGCGACTATCAATGCATG
GCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTTCGCCACAAATAAACATAGTCAATTCCTATGGGGGGATGC
ACAAGTATTTAATCAAGGTTATGGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGG
CAAAGGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAGTGATGAAAGCTATCT
CTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACGATTATTACCAATAGATATTACACCAAAATCTGG
TAGTAGAAAAATGAATGGTAGTAAGGTCACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAA
GTCTGAATTATTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACGGTAAAGATTTTTA
TGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGTATTGAGAAATACAAAGATTGTTGAAGA
CATGGAAAAAGTAAAAGCAACAGAGAGGTTCTTACCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAG
GCACCAAAAAAACATTTGATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCAGTT
GTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTATGGTGTGAAGGAGTTAGAAAT
TGAGAGCATTGCTTTAGGATTAGGTGCTAATAGCAAAGAAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAA
TTGGGAGAGACCCGATACCAAAACCTTTTTAAAAGACTCCTATTATAGTATTAAG

SEQ ID NO. 2811: SAG1552 FROM THE M781 GBS TYPE III STRAIN
TTTGATGGTAGTTTGTATTTACCACAGGGCTTATTAAAAGAAAATACAAGAACTAACTTTGTTGTTAAAGGTGATACT
GTACTTCACAAGCCCACCAATAAACCTTTTGTTGTTAAAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCAC
AACGATTTTCCTATTACTCAAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGGCAAATACTGTAAGA
GTCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGAATCAAAGAGGCCACTGTATTTG
TTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCTATAACAGCTTTTAATGATAATTATAGGGGGTATTTA
AAACGAGAAGCAAAAGGCGTTGTGGATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTTGGTAGCCGTCAT
TATCATTATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGTACTGTCGCTTATACT
AATCATCAAGAGAAAAAAACGCAATATAAAGGACGTTATTTTAAAACTTCTGTGGCAGCTAATCCATTTGAGGTCATG
CTAGCTCAAGTAATGGATGAATTGACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAAC
TCACCAACAACAGACCCTTTTCATTATCGAAAACCATTTGAGGCACAGGCTCCTAAATACGTACAACTAAATGTAGAA
AATATTCAAGCTAATTCAAATGTTAAAGCAGGTATGTTTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAG
GATTATCTATTATTTGATAAAGAGAATATCAGTAAAGAAGATAGACAAAAGATTAAAGAACTTTCTTTGTCACAGGGA
TACGTTAAACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTATTCGACAGCGAGAGGTATT
GCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGAAAAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAA
TCTTTTATATCATCCGGTAGTTTTGGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACA
TCTTTCGCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTATGGTTTATTAGGC
TTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAAAGGAGAGTGGAAACATCCTCTGATGACTAGT
GCAACAGGAGATGACTTATATGCTAGCAGTGATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTA
AAAGAAAAACGATTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGTCACATTT
TCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATTATTTGTCCAAGAGCGCTATAATGCC
TTAAAAGCGAACTATCTTCGACAGCTTAACGGTAAAGATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTT
GAGCAGATAAATATGGTATTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTTA
CCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAAACATTTGATTCACAAACAGATATT
TCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCAGTTGTTGAATTTTTCTGATCCATCATCTCAAAAAATT
CACGATGATTACTTTAAACATTATGGTGTGAAGGAGTTAGAAATTGAGAGCATTGCTTTAGGATTAGGTGCTAATAGC
AAAGAAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACCAAAACCTTTTTAAAA
GACTCCTATTATAGTATTAAGAAAGAATGG

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

```
SEQ2801        -----------------------------------------------------------
SEQ2802        -----------------------------------------------------------
SEQ2803        -----------------------------AAGGGCTTATTAAAAGAAAATACAAGAACT
SEQ2804        ------------------------------------TATTAAAAGAAAATACAAGAACT
SEQ2805        -----------------------------AAGGGCTTATTAAAAGAAAATACAAGAACT
SEQ2806        ATTACTTTGATGGTAGTTTGTATTTACCAAAGGGCTTATTAAAAGAAAATACAAGAACT
SEQ2807        ----------------------TTTACCACAGGGCTTATTAAAAGAAAATACAAGAACT
SEQ2808        ----------------------------AAGGGGCTTATTAAAAGAAAATACAAGAACT
SEQ2809        -----------------------------------------------------------
SEQ2810        -------------------------------------------------TACAAGAACT
SEQ2811        -----TTTGATGGTAGTTTGTATTTACCACAGGGCTTATTAAAAGAAAATACAAGAACT

SEQ2801        --TTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTT
SEQ2802        -----------------------------------------------------------
SEQ2803        ACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTT
SEQ2804        ACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTT
SEQ2805        ACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTT
SEQ2806        ACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTT
SEQ2807        ACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTT
SEQ2808        ACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTT
SEQ2809        ACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTT
SEQ2810        ACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTT
SEQ2811        ACTTTGTTGTTAAAGGTGATACTGTACTTCACAAGCCCACCAATAAACCTTTTGTTGTT

SEQ2801        AAGGAGTAGACGTTGAGTCTTCCTTAGCAGGTTATCATCACAACGATTTTCCTATTACT
SEQ2802        -----------------------------------------------------------
SEQ2803        AAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACT
SEQ2804        AAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACT
SEQ2805        AAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACT
SEQ2806        AAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACT
SEQ2807        AAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACT
SEQ2808        AAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACT
SEQ2809        AAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACT
SEQ2810        AAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACT
SEQ2811        AAGGAGTAGACGTTGAGTCTTCCTTAGCGGGTTATCATCACAACGATTTTCCTATTACT

SEQ2801        AAAAAACGTATCGTGAGTGGTTCCATTTAATTTCCAACATGGGGCAAATACTGTAAGA
SEQ2802        ----------------------------------------------------------
SEQ2803        AAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGCAAATACTGTAAGA
SEQ2804        AAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGCAAATACTGTAAGA
SEQ2805        AAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGCAAATACTGTAAGA
SEQ2806        AAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGCAAATACTGTAAGA
SEQ2807        AAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGCAAATACTGTAAGA
SEQ2808        AAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGCAAATACTGTAAGA
SEQ2809        AAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGCAAATACTGTAAGA
SEQ2810        AAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGCAAATACTGTAAGA
SEQ2811        AAAAAACGTATCGTGAATGGTTCCATTTAATTTCCAACATGGGGCAAATACTGTAAGA

SEQ2801        TCAAAGTACCGATGAATGTTGCATTTTACGATGCTTTATATCACCACAACAAAGCATCA
SEQ2802        -----------------------------------------------------------
SEQ2803        TCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGCATCA
SEQ2804        TCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGCATCA
SEQ2805        TCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGCATCA
SEQ2806        TCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGCATCA
SEQ2807        TCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGAATCA
SEQ2808        TCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGCATCA
SEQ2809        TCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGCATCA
SEQ2810        TCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGAATCA
SEQ2811        TCAAGGTACCGATGAATGTTGCATTTTACGATGCCTTATATCACCACAACAAAGAATCA

SEQ2801        AGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCT
SEQ2802        -----------------------------------------------------------
SEQ2803        AGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCT
SEQ2804        AGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCT
SEQ2805        AGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCT
SEQ2806        AGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCT
SEQ2807        AGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCT
SEQ2808        AGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCT
```

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

```
SEQ2809          AGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCT
SEQ2810          AGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCT
SEQ2811          AGAGGCCACTGTATTTGTTGCAAGGAATACGTATAGATTCTTATCGCAATAATGCTTCT

SEQ2801          TAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
SEQ2802          -----------------------------------------------------------
SEQ2803          TAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
SEQ2804          TAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
SEQ2805          TAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
SEQ2806          TAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
SEQ2807          TAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
SEQ2808          TAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
SEQ2809          TAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
SEQ2810          TAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG
SEQ2811          TAACAGCTTTTAATGATAATTATAGGGGGTATTTAAAACGAGAAGCAAAAGGCGTTGTG

SEQ2801          ATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTTGGTAGCCGTCATTATCAT
SEQ2802          ---------ATGACTA-GTGCAACAGGAGATGACTTATAT-GCTAGCAGTGATGAAAGC
SEQ2803          ATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTGGGTAGCCGTCATTATCAT
SEQ2804          ATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTGGGTAGCCGTCATTATCAT
SEQ2805          ATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTGGGTAGCCGTCATTATCAT
SEQ2806          ATATTCTCCATGGGCGTAAGCAAGTATGGAATACAGATTTTGGTAGCCGTCATTATCAT
SEQ2807          ATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTTGGTAGCCGTCATTATCAT
SEQ2808          ATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTTGGTAGCAGTCATTATCAT
SEQ2809          ATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTTGGTAGCAGTCATTATCAT
SEQ2810          ATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTTGGTAGCCGTCATTATCAT
SEQ2811          ATATTCTCCATGGGCGTAAGCAAGTATGGAATACTGATTTTGGTAGCCGTCATTATCAT

SEQ2801          TATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGT-AC
SEQ2802          TAT--CTCTA--CCTTGCG-ATTAAAACAAAACCTGAAAAACTAAAAGAAAAACGATTAT
SEQ2803          TATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGT-AC
SEQ2804          TATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGT-AC
SEQ2805          TATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGT-AC
SEQ2806          TATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGT-AC
SEQ2807          TATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGT-AC
SEQ2808          TATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATGGACATAGTGGT-AC
SEQ2809          TATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGT-AC
SEQ2810          TATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGCAATAGTGGT-AC
SEQ2811          TATGATCTTAGTCCTTGGGTACTTGGTTATGTCGTAGGGGATGATTGGAATAGTGGT-AC

SEQ2801          TGTCGCTT-ATACTAATCATCAAGAGAA-AAAAACGCAATATAAAGGAC-GTTATTTTAA
SEQ2802          TACCAATAGATATTA--CACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGTCAC
SEQ2803          TGTCGCTT-ATACTAATCATCAAGAGAA-AAAAACGCAATATAAAGGAC-GTTATTTTAA
SEQ2804          TGTCGCTT-ATACTAATCATCAAGAGAA-AAAAACGCAATATAAAGGAC-GTTATTTTAA
SEQ2805          TGTCGCTT-ATACTAATCATCAAGAGAA-AAAAACGCAATATAAAGGAC-GTTATTTTAA
SEQ2806          TGTCGCTT-ATACTAATCATCAAGAGAA-AAAAACGCAATATAAAGGAC-GTTATTTTAA
SEQ2807          TGTCGCTT-ATACTAATCATCAAGAGAA-AAAAACGCAATATAAAGGAC-GTTATTTTAA
SEQ2808          TGTCGCTTTATACTAATCATCAAGAGGAGAAAAACGCAATATAAAGGAC-GTTATTTTAA
SEQ2809          TGTCGCTT-ATACTAATCATCAAGAGAA-AAAAACGCAATATAAAGGAC-GTTATTTTAA
SEQ2810          TGTCGCTT-ATACTAATCATCAAGAGAA-AAAAACGCAATATAAAGGAC-GTTATTTTAA
SEQ2811          TGTCGCTT-ATACTAATCATCAAGAGAA-AAAAACGCAATATAAAGGAC-GTTATTTTAA

SEQ2801          AACTTCTGCGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTTATGGAT--GAATTG
SEQ2802          ATTTTCTAAATCTAGTGA-CTTTGTATTGTC-TATTGATCCAAATGGCAAGTCTGAATT-
SEQ2803          AACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGAT--GAATTG
SEQ2804          AACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGAT--GAATTG
SEQ2805          AACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGAT--GAATTG
SEQ2806          AACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGAT--GAATTG
SEQ2807          AACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGAT--GAATTG
SEQ2808          AACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGAT--GAATTG
SEQ2809          AACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGAT--GAATTG
SEQ2810          AACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGAT--GAATTG
SEQ2811          AACTTCTGTGGCAGCTAATCCATTTGAGGTCATGCTAGCTCAAGTAATGGAT--GAATTG

SEQ2801          ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCA
SEQ2802          ATTTGTC-CAAGAGCGCTATA-ATGCCT--TAAAAGCGAACTATCTTCGACAGCTTAACG
SEQ2803          ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCA
SEQ2804          ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCA
```

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

```
SEQ2805        ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCA
SEQ2806        ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCA
SEQ2807        ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCA
SEQ2808        ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCA
SEQ2809        ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCA
SEQ2810        ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCA
SEQ2811        ACACATTATGAGACAGCTAAATATGGTTGGCAACATTTGATTAGTTTTTCAAACTCACCA

SEQ2801        CAACAGAC------CCTTTTCGTTATCGAAAACCATTTG-AGGCACAGGCTCCTAAATA
SEQ2802        TAAAGATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATCAATA
SEQ2803        CAACAGAC------CCTTTTCATTATCGAAAACCATTTG-AGGCACAGGCTCCTAAATA
SEQ2804        CAACAGAC------CCTTTTCATTATCGAAAACCATTTG-AGGCACAGGCTCCTAAATA
SEQ2805        CAACAGAC------CCTTTTCATTATCGAAAACCATTTG-AGGCACAGGCTCCTAAATA
SEQ2806        CAACAGAC------CCTTTTCATTATCGAAAACCATTTG-AGGCACAGGCTCCTAAATA
SEQ2807        CAACAGAC------CCTTTTCATTATCGAAAACCATTTG-AGGCACAGGCTCCTAAATA
SEQ2808        CAACAGAC------CCTTTTCATTATCGAAAACCATTTG-AGGCACAGGCTCCTAAATA
SEQ2809        CAACAGAC------CCTTTTCATTATCGAAAACCATTTG-AGGCACAGGCTCCTAAATA
SEQ2810        CAACAGAC------CCTTTTCATTATCGAAAACCATTTG-AGGCACAGGCTCCTAAATA
SEQ2811        CAACAGAC------CCTTTTCATTATCGAAAACCATTTG-AGGCACAGGCTCCTAAATA

SEQ2801        G-TACAACTAAATGTAGAAAATATTCAAGCTAATTCGAATGTTAAAGCA---GGTATTT
SEQ2802        GGTATTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAGCAACAGAGAGGT
SEQ2803        G-TACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTTAAAGCA---GGTATGT
SEQ2804        G-TACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTTAAAGCA---GGTATGT
SEQ2805        G-TACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTTAAAGCA---GGTATGT
SEQ2806        G-TACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTTAAAGCA---GGTATGT
SEQ2807        G-TACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTTAAAGCA---GGTATGT
SEQ2808        G-TACAACTAAATGTAGAAAATATTCAAGCTAATTCGAATGTTAAAGCA---GGTATGT
SEQ2809        G-TACAACTAAATGTAGAAAATATTCAAGCTAATTCGAATGTTAAAGCA---GGTATGT
SEQ2810        G-TACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTTAAAGCA---GGTATGT
SEQ2811        G-TACAACTAAATGTAGAAAATATTCAAGCTAATTCAAATGTTAAAGCA---GGTATGT

SEQ2801        TTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATA
SEQ2802        TCTTACCAACTCATCCTACTGGTCTT----CTCAAAACAGGAACAATTGAT-AGGCACCA
SEQ2803        TTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATA
SEQ2804        TTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATA
SEQ2805        TTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATA
SEQ2806        TTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATA
SEQ2807        TTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATA
SEQ2808        TTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATA
SEQ2809        TTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATA
SEQ2810        TTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATA
SEQ2811        TTGCAGCATATAAAGCTATTGATTTCCATCCTCGATACAAGGATTATCTATTATTTGATA

SEQ2801        AAGAGAATATCAGTAAAGAAGATAGACAAAAGATT-AAAGAACTTTCTTTGTCACAGGGA
SEQ2802        AAAAACATTTGATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAAT
SEQ2803        AAGAGAATATCAGTAAAGAAGATAGACAAAAGATT-AAAGAACTTTCTTTGTCACAGGGA
SEQ2804        AAGAGAATATCAGTAAAGAAGATAGACAAAAGATT-AAAGAACTTTCTTTGTCACAGGGA
SEQ2805        AAGAGAATATCAGTAAAGAAGATAGACAAAAGATT-AAAGAACTTTCTTTGTCACAGGGA
SEQ2806        AAGAGAATATCAGTAAAGAAGATAGACAAAAGATT-AAAGAACTTTCTTTGTCACAGGGA
SEQ2807        AAGAGAATATCAGTAAAGAAGATAGACAAAAGATT-AAAGAACTTTCTTTGTCACAGGGA
SEQ2808        AAGAGAATATCAGTAAAGAAGATAGACAAAAGATT-AAAGAACTTTCTTTGTCACAGGGA
SEQ2809        AAGAGAATATCAGTAAAGAAGATAGACAAAAGATT-AAAGAACTTTCTTTGTCACAGGGA
SEQ2810        AAGAGAATATCAGTAAAGAAGATAGACAAAAGATT-AAAGAACTTTCTTTGTCACAGGGA
SEQ2811        AAGAGAATATCAGTAAAGAAGATAGACAAAAGATT-AAAGAACTTTCTTTGTCACAGGGA

SEQ2801        TACGTTA-AACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTA
SEQ2802        TCCGTGGCAGTTGTTGAATTTTTCTGATCCA---TCATCTCAAAAAATTCACGATGATTA
SEQ2803        TACGTTA-AACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTA
SEQ2804        TACGTTA-AACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTA
SEQ2805        TACGTTA-AACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTA
SEQ2806        TACGTTA-AACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTA
SEQ2807        TACGTTA-AACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTA
SEQ2808        TACGTTA-AACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTA
SEQ2809        TACGTTA-AACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTA
SEQ2810        TACGTTA-AACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTA
SEQ2811        TACGTTA-AACTGCTAAATGCTTATCACAAAATCCCTGTTCTAGTCACGGGTTATGGCTA
```

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

```
SEQ2801         TCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGA
SEQ2802         TTTAAACATTATGGTGTGAAGGAGTTAGAAATTGA-GAGCATTGCTTTAGGATTAGGTG
SEQ2803         TCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGA
SEQ2804         TCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGA
SEQ2805         TCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGA
SEQ2806         TCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGA
SEQ2807         TCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGA
SEQ2808         TCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGA
SEQ2809         TCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGA
SEQ2810         TCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGA
SEQ2811         TCGACAGCGAGAGGTATTGCCCAAAAAGAAATTGATAAACGTCCTCTGCCGATTAATGA

SEQ2801         AAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTT
SEQ2802         TAATAGCAAAGAAAACACACTGATAAAGATGGCAGAT------TATCGTTTGAAAAATT
SEQ2803         AAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTT
SEQ2804         AAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTT
SEQ2805         AAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTT
SEQ2806         AAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTT
SEQ2807         AAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTT
SEQ2808         AAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTT
SEQ2809         AAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTT
SEQ2810         AAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTT
SEQ2811         AAAGAACAAGGTCAGCGTTTACTAGAAGATTATGAATCTTTTATATCATCCGGTAGTTT

SEQ2801         GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCCTT
SEQ2802         GGAGAGAC--CCGATAC----CAAAACCTTTTTAA----AAGACTCCTATTATAGTATT
SEQ2803         GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTT
SEQ2804         GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTT
SEQ2805         GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTT
SEQ2806         GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTT
SEQ2807         GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTT
SEQ2808         GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGTGTGGAATACATCCTT
SEQ2809         GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGTGTGGAATACATCCTT
SEQ2810         GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTT
SEQ2811         GGAGCGACTATCAATGCATGGCAAGACGATTGGAATGCAAGGGCGTGGAATACATCTTT

SEQ2801         GCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTA
SEQ2802         A--AGAAAGAA------------------------------------------------
SEQ2803         GCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTA
SEQ2804         GCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTA
SEQ2805         GCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTA
SEQ2806         GCCACAAATAAACATAATCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTA
SEQ2807         GCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTA
SEQ2808         GCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTA
SEQ2809         GCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTA
SEQ2810         GCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTA
SEQ2811         GCCACAAATAAACATAGTCAATTCCTATGGGGGGATGCACAAGTATTTAATCAAGGTTA

SEQ2801         GGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAA
SEQ2802         -----------------------------------------------------------
SEQ2803         GGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAA
SEQ2804         GGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAA
SEQ2805         GGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAA
SEQ2806         GGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAA
SEQ2807         GGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAA
SEQ2808         GGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAGGTTGATGGTAAAAGAGGCAA
SEQ2809         GGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAGGTTGATGGTAAAAGAGGCAA
SEQ2810         GGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAA
SEQ2811         GGTTTATTAGGCTTTAAAAACGCAAAACATCATTATCAAGTTGATGGTAAAAGAGGCAA

SEQ2801         GGAGAGTGGAAACATCCTCTG--------------------------------------
SEQ2802         -----------------------------------------------------------
SEQ2803         GGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAG
SEQ2804         GGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAG
SEQ2805         GGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAG
SEQ2806         GGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAG
SEQ2807         GGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAG
SEQ2808         GAAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAG
```

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

```
SEQ2809          GAAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAG
SEQ2810          GGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAG
SEQ2811          GGAGAGTGGAAACATCCTCTGATGACTAGTGCAACAGGAGATGACTTATATGCTAGCAG

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          GATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACG
SEQ2804          GATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACG
SEQ2805          GATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACG
SEQ2806          GATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACG
SEQ2807          GATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACG
SEQ2808          GATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACG
SEQ2809          GATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACG
SEQ2810          GATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACG
SEQ2811          GATGAAAGCTATCTCTACCTTGCGATTAAAACAAAACCTGAAAAACTAAAAGAAAAACG

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          TTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGT
SEQ2804          TTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGT
SEQ2805          TTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGT
SEQ2806          TTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGT
SEQ2807          TTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGT
SEQ2808          TTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGT
SEQ2809          TTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGT
SEQ2810          TTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGT
SEQ2811          TTATTACCAATAGATATTACACCAAAATCTGGTAGTAGAAAAATGAATGGTAGTAAGGT

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          ACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATT
SEQ2804          ACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATT
SEQ2805          ACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATT
SEQ2806          ACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATT
SEQ2807          ACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATT
SEQ2808          ACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATT
SEQ2809          ACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATT
SEQ2810          ACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATT
SEQ2811          ACATTTTCTAAATCTAGTGACTTTGTATTGTCTATTGATCCAAATGGCAAGTCTGAATT

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          TTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACGGTAA
SEQ2804          TTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACGGTAA
SEQ2805          TTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACGGTAA
SEQ2806          TTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACGGTAA
SEQ2807          TTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACGGTAA
SEQ2808          TTTGTCCAAGAGCGCTATAACGCCTTAAAAGCGAACTATCTTCGACAGCTTAATGGTAA
SEQ2809          TTTGTCCAAGAGCGCTATAACGCCTTAAAAGCGAACTATCTTCGACAGCTTAATGGTAA
SEQ2810          TTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACGGTAA
SEQ2811          TTTGTCCAAGAGCGCTATAATGCCTTAAAAGCGAACTATCTTCGACAGCTTAACGGTAA

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          GATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGT
SEQ2804          GATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGT
SEQ2805          GATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGT
SEQ2806          GATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGT
SEQ2807          GATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGT
SEQ2808          GATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGT
SEQ2809          GATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGT
SEQ2810          GATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGT
SEQ2811          GATTTTTATGCTTTCCCACCAAAGAAGAACAGTAGTAATTTTGAGCAGATAAATATGGT

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          TTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTT
SEQ2804          TTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTT
```

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

```
SEQ2805          TTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTT
SEQ2806          TTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTT
SEQ2807          TTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTT
SEQ2808          TTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTT
SEQ2809          TTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTT
SEQ2810          TTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTT
SEQ2811          TTGAGAAATACAAAGATTGTTGAAGACATGGAAAAAGTAAAAGCAACAGAGAGGTTCTT

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          CCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATT
SEQ2804          CCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATT
SEQ2805          CCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATT
SEQ2806          CCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATT
SEQ2807          CCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATT
SEQ2808          CCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATT
SEQ2809          CCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATT
SEQ2810          CCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATT
SEQ2811          CCAACTCATCCTACTGGTCTTCTCAAAACAGGAACAACTGATAGGCACCAAAAAACATT

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          GATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCA
SEQ2804          GATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCA
SEQ2805          GATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCA
SEQ2806          GATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCA
SEQ2807          GATTCACAACCAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCA
SEQ2808          GATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCA
SEQ2809          GATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCA
SEQ2810          GATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCA
SEQ2811          GATTCACAAACAGATATTTCGTTTGGAAAGGACTTTATAGAGGTCAGAATTCCGTGGCA

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          TTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTA
SEQ2804          TTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTA
SEQ2805          TTGTTGAATTTTTCTGATCCATCATCTCAAAGAATTCACGATGATTACTTTAAACATTA
SEQ2806          TTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTA
SEQ2807          TTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTA
SEQ2808          TTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTA
SEQ2809          TTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTA
SEQ2810          TTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTA
SEQ2811          TTGTTGAATTTTTCTGATCCATCATCTCAAAAAATTCACGATGATTACTTTAAACATTA

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          GGTGTGAAGGAGTTAGAAATTGAGAG--CATTGCTTTAGGATTAGGTGCTAATAGCAAA
SEQ2804          GGTGTGAAGGAGTTAGAAATTGAGAG--CATTGCTTTAGGATTAGGTGCTAATAGCAAA
SEQ2805          GGTGTGAAGGAGTTAGAAAATTGAGAGCCATTGCTTTAGGATTAGGTGCTAATAGCAAA
SEQ2806          GGTGTGAAGGAGTTAGAAATTGAGAG--CATTGCTTTAGGATTAGGTGCTAATAGCAAA
SEQ2807          GGTGTGAAGGAGTTAGAAATTGAGAG--CATTGCTTTAGGATTAGGTGCTAATAGCAAA
SEQ2808          GGTGTGAAGGAGTTAGAAATTGAGAG--CATTGCTTTAGGATTAGGTGCTAATAGCAAA
SEQ2809          GGTGTGAAGGAGTTAGAAATTGAGAG--CATTGCTTTAGGATTAGGTGCTAATAGCAAA
SEQ2810          GGTGTGAAGGAGTTAGAAATTGAGAG--CATTGCTTTAGGATTAGGTGCTAATAGCAAA
SEQ2811          GGTGTGAAGGAGTTAGAAATTGAGAG--CATTGCTTTAGGATTAGGTGCTAATAGCAAA

SEQ2801          ------------------------------------------------------------
SEQ2802          ------------------------------------------------------------
SEQ2803          AAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACC
SEQ2804          AAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACC
SEQ2805          AAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACC
SEQ2806          AAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACC
SEQ2807          AAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACC
SEQ2808          AAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACC
SEQ2809          AAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACC
SEQ2810          AAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACC
SEQ2811          AAAACACACTGATAAAGATGGCAGATTATCGTTTGAAAAATTGGGAGAGACCCGATACC
```

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

```
SEQ2801     -----------------------------------------------------------
SEQ2802     -----------------------------------------------------------
SEQ2803     AAACCTTTTTAAAAGACTCCTATTATGTATTAAGAAAGAA-------------------
SEQ2804     AAACCTTTTTAAAAGACTCCTATTATAGTATTAAGAAAGAATGGTCTAAAGAAAGAGAG
SEQ2805     AAACCTTTTTAAAAGA-------------------------------------------
SEQ2806     AAACCTTTTTAAAAGACTCCTATTATGTATTAAGAAAGA--------------------
SEQ2807     AAACCTTTTTAAAAGACT-----------------------------------------
SEQ2808     AAACCTTTTTAAAAGACTCCTATTATAGT------------------------------
SEQ2809     AAACCTTTTTAAAAGACTCCTATTATAGTATTAAGAAAG--------------------
SEQ2810     AAACCTTTTTAAAAGACTCCTATTATAGTATTAAG------------------------
SEQ2811     AAACCTTTTTAAAAGACTCCTATTATAGTATTAAGAAAGAATGG---------------

SEQ2801     --------------
SEQ2802     --------------
SEQ2803     --------------
SEQ2804     GAACATATGGTCCA
SEQ2805     --------------
SEQ2806     --------------
SEQ2807     --------------
SEQ2808     --------------
SEQ2809     --------------
SEQ2810     --------------
SEQ2811     --------------
```

>SEQ ID NO 2850:62_1169NT frame: 1

```
FVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPITQKTYREWFHLISNMGANTVRV
KVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNASITAFNDNYRGYLKREAKGVVD
ILHGRKQVWNTDFGSRHYHYDLSPWVLGYVVGDDWNSGTVAYTNHQEKKTQYKGRYFKTS
AAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTTDPFRYRKPFEAQAPKYVQLNV
ENIQANSNVKAGIFAAYKAIDFHPRYKDYLLFDKENISKEDRQKIKELSLSQGYVKLLNA
YHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQRLLEDYESFISSGSFGATINAW
QDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFKNAKHHYQVDGKRGKGEWKHPL
MTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDITPKSGSRKMNGSKVTFSKSSD
FVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFPPKKNSSNFEQINMVLRNTKIV
EDMEKVKATERFLPTHPTGLLKTGTIDRHQKTFDSQTDISFGKDFIEVRIPWQLLNFSDP
SSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKMADYRLKNWERPDTKTFLKDSY
YSI.ER
```

>SEQ ID NO 2851:62_18RS21 frame: 1

```
KGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPITQKTYREWFHL
ISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNASITAFNDNYRG
YLKREAKGVVDILHGRKQVWNTDLGSRHYHYDLSPWVLGYVVGDDWNSGTVAYTNHQEKK
TQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTTDPFHYRKPFE
AQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISKEDRQKIKELS
LSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQRLLEDYESFIS
SGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFKNAKHHYQVDG
KRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDITPKSGSRKMN
GSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFPPKKNSSNFEQ
INMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDISFGKDFIEVR
IPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKMADYRLKNWER
PDTKTFLKDSYYVLRK
```

>SEQ ID NO 2852:62_2603 frame: 3

```
LKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPITQKTYREWFHLISN
MGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNASITAFNDNYRGYLK
REAKGVVDILHGRKQVWNTDLGSRHYHYDLSPWVLGYVVGDDWNSGTVAYTNHQEKKTQY
KGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTTDPFHYRKPFEAQA
PKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISKEDRQKIKELSLSQ
GYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQRLLEDYESFISSGS
FGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFKNAKHHYQVDGKRG
KGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDITPKSGSRKMNGSK
VTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFPPKKNSSNFEQINM
VLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDISFGKDFIEVRIPW
QLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKMADYRLKNWERPDT
KTFLKDSYYSIKKEWSKERERTYGP
```

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

>SEQ ID NO 2853:62_A909 frame: 1
KGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPITQKTYREWFHL
ISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNASITAFNDNYRG
YLKREAKGVVDILHGRKQVWNTDLGSRHYHYDLSPWVLGYVVGDDWNSGTVAYTNHQEKK
TQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTTDPFHYRKPFE
AQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISKEDRQKIKELS
LSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQRLLEDYESFIS
SGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFKNAKHHYQVDG
KRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDITPKSGSRKMN
GSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFPPKKNSSNFEQ
INMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDISFGKDFIEVR
IPWQLLNFSDPSSQRIHDDYFKHYGVKELEN.EPLL.D.VLIAKKTH..RWQIIV.KIGR
DPIPKPF.K

>SEQ ID NO 2854:62_A909 frame: 1
KGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPITQKTYREWFHL
ISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNASITAFNDNYRG
YLKREAKGVVDILHGRKQVWNTDLGSRHYHYDLSPWVLGYVVGDDWNSGTVAYTNHQEKK
TQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTTDPFHYRKPFE
AQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISKEDRQKIKELS
LSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQRLLEDYESFIS
SGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFKNAKHHYQVDG
KRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDITPKSGSRKMN
GSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFPPKKNSSNFEQ
INMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDISFGKDFIEVR
IPWQLLNFSDPSSQRIHDDYFKHYGVKELEN.EPLL.D.VLIAKKTH..RWQIIV.KIGR
DPIPKPF.K

>SEQ ID NO 2855:62_CJB110 frame: 1
YYFDGSLYLPKGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNAS
ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDFGSRHYHYDLSPWVLGYVVGDDWNSGT
VAYTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT
DPFHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISK
EDRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR
LLEDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHNQFLWGDAQVFNQGYGLLGFK
NAKHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI
TPKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP
PKKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDI
SFGKDFIEVRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKM
ADYRLKNWERPDTKTFLKDSYYVLRK

>SEQ ID NO 2856:62_COH1 frame: 2
LPQGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPITQKTYREWF
HLISNMGANTVRVKVPMNVAFYDALYHHNKESKRPLYLLQGIRIDSYRNNASITAFNDNY
RGYLKREAKGVVDILHGRKQVWNTDFGSRHYHYDLSPWVLGYVVGDDWNSGTVAYTNHQE
KKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTTDPFHYRKP
FEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISKEDRQKIKE
LSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQRLLEDYESF
ISSGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFKNAKHHYQV
DGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDITPKSGSRK
MNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFPPKKNSSNF
EQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQPDISFGKDFIE
VRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKMADYRLKNW
ERPDTKTFLKD

>SEQ ID NO 2857:62_H36B frame: 2
RGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPITQKTYREWFHL
ISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNASITAFNDNYRG
YLKREAKGVVDILHGRKQVWNTDFGSSHYHYDLSPWVLGYVVGDDGHSGTVALY

>SEQ ID NO 2858:62_JM9130013 frame: 3
FVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPITQKTYREWFHLISNMGANTVRV
KVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNASITAFNDNYRGYLKREAKGVVD
ILHGRKQVWNTDFGSSHYHYDLSPWVLGYVVGDDWNSGTVAYTNHQEKKTQYKGRYFKTS
VAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTTDPFHYRKPFEAQAPKYVQLNV
ENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISKEDRQKIKELSLSQGYVKLLNA
YHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQRLLEDYESFISSGSFGATINAW

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

```
QDDWNARVWNTSFATNKHSQFLWGDAQVFNQGYGLLGFKNAKHHYQVDGKRGKEEWKHPL
MTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDITPKSGSRKMNGSKVTFSKSSD
FVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFPPKKNSSNFEQINMVLRNTKIV
EDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDISFGKDFIEVRIPWQLLNFSDP
SSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKMADYRLKNWERPDTKTFLKDSY
YSIKK
```

>SEQ ID NO 2859:62_M732 frame: 2

```
TRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPITQKTYREWFHLISNMGAN
TVRVKVPMNVAFYDALYHHNKESKRPLYLLQGIRIDSYRNNASITAFNDNYRGYLKREAK
GVVDILHGRKQVWNTDFGSRHYHYDLSPWVLGYVVGDDCNSGTVAYTNHQEKKTQYKGRY
FKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTTDPFHYRKPFEAQAPKYV
QLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISKEDRQKIKELSLSQGYVK
LLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQRLLEDYESFISSGSFGAT
INAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFKNAKHHYQVDGKRGKGEW
KHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDITPKSGSRKMNGSKVTFS
KSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFPPKKNSSNFEQINMVLRN
TKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDISFGKDFIEVRIPWQLLN
FSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKMADYRLKNWERPDTKTFL
KDSYYSIK
```

>SEQ ID NO 2860:62_M781 frame: 1

```
FDGSLYLPQGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPITQK
TYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKESKRPLYLLQGIRIDSYRNNASIT
AFNDNYRGYLKREAKGVVDILHGRKQVWNTDFGSRHYHYDLSPWVLGYVVGDDWNSGTVA
YTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTTDP
FHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISKED
RQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQRLL
EDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFKNA
KHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDITP
KSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFPPK
KNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDISF
GKDFIEVRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKMAD
YRLKNWERPDTKTFLKDSYYSIKKEW
```

```
SEQ2850         --------------------FVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
SEQ2851         ---------KGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
SEQ2852         ------------LKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
SEQ2853         ---------KGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
SEQ2854         ---------KGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
SEQ2855         YFDGSLYLPKGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
SEQ2856         -------LPQGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
SEQ2857         ---------RGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
SEQ2858         --------------------FVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
SEQ2859         ----------------TRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT
SEQ2860         -FDGSLYLPQGLLKENTRTNFVVKGDTVLHKPTNKPFVVKGVDVESSLAGYHHNDFPIT

SEQ2850         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNAS
SEQ2851         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNAS
SEQ2852         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNAS
SEQ2853         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNAS
SEQ2854         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNAS
SEQ2855         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNAS
SEQ2856         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKESKRPLYLLQGIRIDSYRNNAS
SEQ2857         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNAS
SEQ2858         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKASKRPLYLLQGIRIDSYRNNAS
SEQ2859         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKESKRPLYLLQGIRIDSYRNNAS
SEQ2860         QKTYREWFHLISNMGANTVRVKVPMNVAFYDALYHHNKESKRPLYLLQGIRIDSYRNNAS

SEQ2850         ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDFGSRHYHYDLSPWVLGYVVGDDWNSGT
SEQ2851         ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDLGSRHYHYDLSPWVLGYVVGDDWNSGT
SEQ2852         ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDLGSRHYHYDLSPWVLGYVVGDDWNSGT
SEQ2853         ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDLGSRHYHYDLSPWVLGYVVGDDWNSGT
SEQ2854         ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDLGSRHYHYDLSPWVLGYVVGDDWNSGT
SEQ2855         ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDFGSRHYHYDLSPWVLGYVVGDDWNSGT
SEQ2856         ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDFGSRHYHYDLSPWVLGYVVGDDWNSGT
SEQ2857         ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDFGSSHYHYDLSPWVLGYVVGDDGHSGT
SEQ2858         ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDFGSSHYHYDLSPWVLGYVVGDDWNSGT
```

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

```
SEQ2859          ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDFGSRHYHYDLSPWVLGYVVGDDCNSGT
SEQ2860          ITAFNDNYRGYLKREAKGVVDILHGRKQVWNTDFGSRHYHYDLSPWVLGYVVGDDWNSGT
SEQ2850          VAYTNHQEKKTQYKGRYFKTSAAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT
SEQ2851          VAYTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT
SEQ2852          VAYTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT
SEQ2853          VAYTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT
SEQ2854          VAYTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT
SEQ2855          VAYTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT
SEQ2856          VAYTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT
SEQ2857          VALY--------------------------------------------------------
SEQ2858          VAYTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT
SEQ2859          VAYTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT
SEQ2860          VAYTNHQEKKTQYKGRYFKTSVAANPFEVMLAQVMDELTHYETAKYGWQHLISFSNSPTT

SEQ2850          PFRYRKPFEAQAPKYVQLNVENIQANSNVKAGIFAAYKAIDFHPRYKDYLLFDKENISK
SEQ2851          PFHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISK
SEQ2852          PFHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISK
SEQ2853          PFHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISK
SEQ2854          PFHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISK
SEQ2855          PFHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISK
SEQ2856          PFHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISK
SEQ2857          -----------------------------------------------------------
SEQ2858          PFHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISK
SEQ2859          PFHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISK
SEQ2860          PFHYRKPFEAQAPKYVQLNVENIQANSNVKAGMFAAYKAIDFHPRYKDYLLFDKENISK

SEQ2850          DRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR
SEQ2851          DRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR
SEQ2852          DRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR
SEQ2853          DRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR
SEQ2854          DRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR
SEQ2855          DRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR
SEQ2856          DRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR
SEQ2857          -----------------------------------------------------------
SEQ2858          DRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR
SEQ2859          DRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR
SEQ2860          DRQKIKELSLSQGYVKLLNAYHKIPVLVTGYGYSTARGIAQKEIDKRPLPINEKEQGQR

SEQ2850          LEDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFK
SEQ2851          LEDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFK
SEQ2852          LEDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFK
SEQ2853          LEDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFK
SEQ2854          LEDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFK
SEQ2855          LEDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHNQFLWGDAQVFNQGYGLLGFK
SEQ2856          LEDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFK
SEQ2857          -----------------------------------------------------------
SEQ2858          LEDYESFISSGSFGATINAWQDDWNARVWNTSFATNKHSQFLWGDAQVFNQGYGLLGFK
SEQ2859          LEDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFK
SEQ2860          LEDYESFISSGSFGATINAWQDDWNARAWNTSFATNKHSQFLWGDAQVFNQGYGLLGFK

SEQ2850          AKHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI
SEQ2851          AKHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI
SEQ2852          AKHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI
SEQ2853          AKHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI
SEQ2854          AKHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI
SEQ2855          AKHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI
SEQ2856          AKHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI
SEQ2857          -----------------------------------------------------------
SEQ2858          AKHHYQVDGKRGKEEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI
SEQ2859          AKHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI
SEQ2860          AKHHYQVDGKRGKGEWKHPLMTSATGDDLYASSDESYLYLAIKTKPEKLKEKRLLPIDI

SEQ2850          PKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP
SEQ2851          PKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP
SEQ2852          PKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP
SEQ2853          PKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP
SEQ2854          PKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP
SEQ2855          PKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP
```

Table 28: Comparative Sequences relating to SAG1552 (conserved hypothetical protein)

```
SEQ2856         PKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP
SEQ2857         ------------------------------------------------------------
SEQ2858         PKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP
SEQ2859         PKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP
SEQ2860         PKSGSRKMNGSKVTFSKSSDFVLSIDPNGKSELFVQERYNALKANYLRQLNGKDFYAFP

SEQ2850         KKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTIDRHQKTFDSQTDI
SEQ2851         KKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDI
SEQ2852         KKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDI
SEQ2853         KKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDI
SEQ2854         KKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDI
SEQ2855         KKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDI
SEQ2856         KKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQPDI
SEQ2857         ------------------------------------------------------------
SEQ2858         KKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDI
SEQ2859         KKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDI
SEQ2860         KKNSSNFEQINMVLRNTKIVEDMEKVKATERFLPTHPTGLLKTGTTDRHQKTFDSQTDI

SEQ2850         FGKDFIEVRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKM
SEQ2851         FGKDFIEVRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKM
SEQ2852         FGKDFIEVRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKM
SEQ2853         FGKDFIEVRIPWQLLNFSDPSSQRIHDDYFKHYGVKELENEPLLDVLIAKKTHRWQIIV
SEQ2854         FGKDFIEVRIPWQLLNFSDPSSQRIHDDYFKHYGVKELENEPLLDVLIAKKTHRWQIIV
SEQ2855         FGKDFIEVRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKM
SEQ2856         FGKDFIEVRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKM
SEQ2857         ------------------------------------------------------------
SEQ2858         FGKDFIEVRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKM
SEQ2859         FGKDFIEVRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKM
SEQ2860         FGKDFIEVRIPWQLLNFSDPSSQKIHDDYFKHYGVKELEIESIALGLGANSKENTLIKM

SEQ2850         DYRLKNWERPDTKTFLKDSYYSIER------------
SEQ2851         DYRLKNWERPDTKTFLKDSYYVLRK------------
SEQ2852         DYRLKNWERPDTKTFLKDSYYSIKKEWSKERERTYGP
SEQ2853         IGRDPIPKPFK--------------------------
SEQ2854         IGRDPIPKPFK--------------------------
SEQ2855         DYRLKNWERPDTKTFLKDSYYVLRK------------
SEQ2856         DYRLKNWERPDTKTFLKD-------------------
SEQ2857         -------------------------------------
SEQ2858         DYRLKNWERPDTKTFLKDSYYSIKK------------
SEQ2859         DYRLKNWERPDTKTFLKDSYYSIK-------------
SEQ2860         DYRLKNWERPDTKTFLKDSYYSIKKEW----------
```

Table 29: Comparative Sequences relating to SAG1641 (YaeC family protein)

SEQ ID NO. 2901: SAG1641 FROM THE 090 GBS TYPE Ia STRAIN
AATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAA
GCACGTTGGGATAAAATTGAAAAGCTAGTAGGCGATAAAGCTAAAATCAAATTCACAGAATTTACAGATTATACACAA
CCAAATCAAGCGACAGCCAATAAGGATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAG
GAAAATAAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCCCCAATTCGTATCTATTCTGAGAAGGTAAAATCT
CTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTCCTT
CAGTCAGCAGGTTTAATCAAATTGAATGTTTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAA
GATATTAATATTCAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAAT
ACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGG
ATTAATATCATTGCGGGACGTAAAAATTGGAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTAT
CACACAGATGAAGTGAAAAAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGGAACCCAGCTTTCTTGTACAA

SEQ ID NO. 2902: SAG1641 FROM THE 1169NT1 GBS TYPE V STRAIN
(REVERSE COMPLEMENT)
ATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAAG
CACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCTAAAATCAAATTTACAGAATTTACAGATTATACACAAC
CAAATCAAGCGACAGCCAATAAGGATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGG
AAAATAAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAGAAGGTAAAATCTC
TTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTCCTTC
AGTCAGCAGGTTTAATCAAATTGAATGTTTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGG
ATATTAATATTCAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAATA
CATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGGA
TTAATATCATTGCGGGACGTAAAAATTGGAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATC
ACACAGATGAAGTGAAAAAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG

SEQ ID NO. 2903: SAG1641 FROM THE 18RS21 GBS TYPE II STRAIN
AATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAA
GCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCTAAAATCAAATTTACAGAATTTACAGATTATACACAA
CCAAATCAAGCGACAGCCAATAAGGATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAG
GAAAATAAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAGAAGGTAAAATCT
CTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTCCTT
CAGTCAGCAGGTTTAATCAAATTGAATGTTTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAG
GATATTAATATTCAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAAT
ACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGG
ATTAATATCATTGCGGGACGTAAAAATTGGAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTAT
CACACAGATGAAGTGAAAAAAGTTATCAAAGATACTTCAGCTGATATTCCAC

SEQ ID NO. 2904: SAG1641 FROM THE 2603 V/R GBS TYPE V STRAIN
AATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAA
GCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCTAAAATCAAATTTACAGAATTTACAGATTATACACAA
CCAAATCAAGCGACAGCCAATAAGGATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAG
GAAAATAAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAGAAGGTAAAATCT
CTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTCCTT
CAGTCAGCAGGTTTAATCAAATTGAATGTTTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAG
GATATTAATATTCAGGAGTTAGATGCCAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAAT
ACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGG
ATTAATATCATTGCGGGACGTAAAAATTGGAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTAT
CACACAGATGAAGTGAAAAAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG

SEQ ID NO. 2905: SAG1641 FROM THE A909 GBS TYPE Ia STRAIN
AATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAA
GCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCTAAAATCAAATTTACAGAATTTACAGATTATACACAA
CCAAATCAAGCGACAGCCAATAAGGATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAG
GAAAATAAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAGAAGGTAAAATCT
CTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTCCTT
CAGTCAGCAGGTTTAATCAAATTGAATGTTTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAG
GATATTAATATTCAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAAT
ACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGG
ATTAATATCATTGCGGGACGTAAAAATTGGAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTAT
CACACAGATGAAGTGAAAAAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG

Table 29: Comparative Sequences relating to SAG1641 (YaeC family protein)

SEQ ID NO. 2906: SAG1641 FROM THE CJB110 GBS NONTYPEABLE STRAIN

AAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGT
AGGCGATAAAGCTAAAATCAAATTCACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAGGATGT
GGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAATAAGAAAAACTTAATTCCACTTGA
AAAGACTTACTTAGCCCCAATTCGTATCTATTCTGAGAAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTAT
TGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGT
TTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAAGATATTAATATTCAGGAGTTAGATGCGAG
TCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAATACATACATTGAGCAAGCTAATTTAAAACC
TTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTG
GAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAAAAAGTTATCAA
AGATACTTCAGCTGATATTCCACAATGGAA

SEQ ID NO. 2907: SAG1641 FROM THE COH1 GBS TYPE III STRAIN
(REVERSE COMPLEMENT)

AGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAAGCACGTTG
GGATAAAATTGAAAAGCTAGTAGGTGATAAAGCTAAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCA
AGCGACAGCCAATAAGGATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAATAA
GAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAGAAGGTAAAATCTCTTAAAAA
ATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTACTTCAGTCAGC
AGGTTTAATCAAATTGAATGTTTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAA
TATTCAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAATACATACAT
TGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGGATTAATAT
CATTGCGGGACGTAAAAATTGGAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGA
TGAAGTGAAAAAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG

SEQ ID NO. 2908: SAG1641 FROM THE H36b GBS TYPE Ib STRAIN

AAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAAGCAC
GTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCTAAAATCAAATTTACAGAATTTACAGATTATACACAACCAA
ATCAAGCGACAGCCAATAAGGATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAA
ATAAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAGAAGGTAAAATCTCTTA
AAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTCCTTCAGT
CAGCAGGTTTAATCAAATTGAATGTTTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATA
TTAATATTCAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAATACAT
ACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGGATTA
ATATCATTGCGGGACGTAAAAATTGGAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACA
CAGATGAAGTGAAAAAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG

SEQ ID NO. 2909: SAG1641 FROM THE JM3190013 GBS TYPE VIII STRAIN

TTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAAGCACGTTGGGA
TAAAATTGAAAAGCTAGTAGGTGATAAAGCTAAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCAAGC
GACAGCCAATAAGGATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAATAAGAA
AAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAGAAGGTAAAATCTCTTAAAAAATT
GAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGG
TTTAATCAAATTGAATGTTTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAATAT
TCAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAATACATACATTGA
GCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGGATTAATATCAT
TGCGGGACGTAAAAATTGGAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGA
AGTGAAAAAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG

SEQ ID NO. 2910: SAG1641 FROM THE M732 GBS TYPE III STRAIN

AATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAA
GCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCTAAAATCAAATTTACAGAATTTACAGATTATACACAA
CCAAATCAAGCGACAGCCAATAAGGATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAG
GAAAATAAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAGAAGGTAAAATCT
CTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTCCTT
CAGTCAGCAGGTTTAATCAAATTGAATGTTTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAG
GATATTAATATTCAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAAT
ACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGG
ATTAATATCATTGCGGGACGTAAAAATTGGAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTAT
CACACAGATGAAGTGAAAAAAGTTATCAAAGATAC

Table 29: Comparative Sequences relating to SAG1641 (YaeC family protein)

SEQ ID NO. 2911: SAG1641 FROM THE M781 GBS TYPE III STRAIN
AGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACCTTTTCTGACACTGAAAAAGCACGTTG
GGATAAAATTGAAAAGCTAGTAGGTGATAAAGCTAAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCA
AGCGACAGCCAATAAGGATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAATAA
GAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAGAAGGTAAAATCTCTTAAAAA
ATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCAACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGC
AGGTTTAATCAAATTGAATGTTTCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAA
TATTCAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATTAATAATACATACAT
TGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAATCAGATAAAAATTCAAAACAATGGATTAATAT
CATTGCGGGACGTAAAAATTGGAAAAAGCAAAAGAACGCTAAAGCTATCCAAGCTATCTGGGATGCTTATCACACAGA
TGAAGTGAAAAAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG

```
SEQ2901         ATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACC
SEQ2902         ATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACC
SEQ2903         ATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACC
SEQ2904         ATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACC
SEQ2905         ATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACC
SEQ2906         ----------------------------AAGTAAAGTTGTTAAAGTTGGTGTTATGACC
SEQ2907         -------AGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACC
SEQ2908         ---AAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACC
SEQ2909         ----------TTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACC
SEQ2910         ATCAAGAAGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACC
SEQ2911         -------AGTTTCAGCAAGCTCAACTTCAAGTAAAGTTGTTAAAGTTGGTGTTATGACC

SEQ2901         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGCGATAAAGCT
SEQ2902         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCT
SEQ2903         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCT
SEQ2904         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCT
SEQ2905         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCT
SEQ2906         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGCGATAAAGCT
SEQ2907         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCT
SEQ2908         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCT
SEQ2909         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCT
SEQ2910         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCT
SEQ2911         TTTTCTGACACTGAAAAAGCACGTTGGGATAAAATTGAAAAGCTAGTAGGTGATAAAGCT

SEQ2901         AAAATCAAATTCACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG
SEQ2902         AAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG
SEQ2903         AAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG
SEQ2904         AAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG
SEQ2905         AAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG
SEQ2906         AAAATCAAATTCACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG
SEQ2907         AAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG
SEQ2908         AAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG
SEQ2909         AAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG
SEQ2910         AAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG
SEQ2911         AAAATCAAATTTACAGAATTTACAGATTATACACAACCAAATCAAGCGACAGCCAATAAG

SEQ2901         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT
SEQ2902         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT
SEQ2903         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT
SEQ2904         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT
SEQ2905         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT
SEQ2906         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT
SEQ2907         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT
SEQ2908         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT
SEQ2909         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT
SEQ2910         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT
SEQ2911         GATGTGGATATTAATGCCTTTCAACATTACAATTTCTTAGAAAACTGGAATAAGGAAAAT

SEQ2901         AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCCCCAATTCGTATCTATTCTGAG
SEQ2902         AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAG
SEQ2903         AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAG
SEQ2904         AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAG
SEQ2905         AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAG
SEQ2906         AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCCCCAATTCGTATCTATTCTGAG
SEQ2907         AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAG
SEQ2908         AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAG
```

Table 29: Comparative Sequences relating to SAG1641 (YaeC family protein)

```
SEQ2909        AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAG
SEQ2910        AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAG
SEQ2911        AAGAAAAACTTAATTCCACTTGAAAAGACTTACTTAGCTCCAATTCGTATCTATTCTGAG

SEQ2901        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA
SEQ2902        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA
SEQ2903        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA
SEQ2904        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA
SEQ2905        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA
SEQ2906        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA
SEQ2907        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA
SEQ2908        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA
SEQ2909        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA
SEQ2910        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA
SEQ2911        AAGGTAAAATCTCTTAAAAAATTGAAAAAAGGAGCCACTATTGCAATTCCAAATGATGCA

SEQ2901        ACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGTT
SEQ2902        ACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGTT
SEQ2903        ACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGTT
SEQ2904        ACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGTT
SEQ2905        ACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGTT
SEQ2906        ACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGTT
SEQ2907        ACAAATGGTAGCCGTGCATTGTATGTACTTCAGTCAGCAGGTTTAATCAAATTGAATGTT
SEQ2908        ACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGTT
SEQ2909        ACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGTT
SEQ2910        ACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGTT
SEQ2911        ACAAATGGTAGCCGTGCATTGTATGTCCTTCAGTCAGCAGGTTTAATCAAATTGAATGTT

SEQ2901        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAAGATATTAATATT
SEQ2902        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAATATT
SEQ2903        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAATATT
SEQ2904        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAATATT
SEQ2905        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAATATT
SEQ2906        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAAGATATTAATATT
SEQ2907        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAATATT
SEQ2908        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAATATT
SEQ2909        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAATATT
SEQ2910        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAATATT
SEQ2911        TCTGGTAAGAAGGTTGCAACAGTTGCTAATATCACATCTAATAAAAAGGATATTAATATT

SEQ2901        CAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT
SEQ2902        CAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT
SEQ2903        CAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT
SEQ2904        CAGGAGTTAGATGCCAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT
SEQ2905        CAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT
SEQ2906        CAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT
SEQ2907        CAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT
SEQ2908        CAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT
SEQ2909        CAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT
SEQ2910        CAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT
SEQ2911        CAGGAGTTAGATGCGAGTCAAACACCACGTGCACTCAAAGATGTAGATGCAGCTATTATT

SEQ2901        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
SEQ2902        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
SEQ2903        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
SEQ2904        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
SEQ2905        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
SEQ2906        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
SEQ2907        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
SEQ2908        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
SEQ2909        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
SEQ2910        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
SEQ2911        AATAATACATACATTGAGCAAGCTAATTTAAAACCTTCAGATGCTATCTTTGTTGAGAAA
```

Table 29: Comparative Sequences relating to SAG1641 (YaeC family protein)

```
SEQ2901         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG
SEQ2902         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG
SEQ2903         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG
SEQ2904         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG
SEQ2905         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG
SEQ2906         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG
SEQ2907         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG
SEQ2908         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG
SEQ2909         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG
SEQ2910         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG
SEQ2911         TCAGATAAAAATTCAAAACAATGGATTAATATCATTGCGGGACGTAAAAATTGGAAAAAG

SEQ2901         CAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAA
SEQ2902         CAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAA
SEQ2903         CAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAA
SEQ2904         CAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAA
SEQ2905         CAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAA
SEQ2906         CAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAA
SEQ2907         CAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAA
SEQ2908         CAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAA
SEQ2909         CAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAA
SEQ2910         CAAAAGAACGCTAAAGCTATCCAAGCTATCTTGGATGCTTATCACACAGATGAAGTGAAA
SEQ2911         CAAAAGAACGCTAAAGCTATCCAAGCTATCTGGGATGCTTATCACACAGATGAAGTGAAA

SEQ2901         AAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGGAACCCAGCTTTCTTGTACAA
SEQ2902         AAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG--------------------
SEQ2903         AAAGTTATCAAAGATACTTCAGCTGATATTCCAC-------------------------
SEQ2904         AAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG--------------------
SEQ2905         AAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG--------------------
SEQ2906         AAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGGAA------------------
SEQ2907         AAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG--------------------
SEQ2908         AAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG--------------------
SEQ2909         AAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG--------------------
SEQ2910         AAAGTTATCAAAGATAC------------------------------------------
SEQ2911         AAAGTTATCAAAGATACTTCAGCTGATATTCCACAATGG--------------------
```

```
>SEQ ID NO 2950: 35_090 frame: 1
NQEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
KVIKDTSADIPQWNPAFLY

>SEQ ID NO 2951: 35_1169NT frame: 3
QEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANKD
VDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDAT
NGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAIIN
NTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVKK
VIKDTSADIPQW

>SEQ ID NO 2952: 35_18RS21 frame: 1
NQEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
KVIKDTSADIP

>SEQ ID NO 2953:35_2603 frame: 1
NQEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
KVIKDTSADIPQW
```

Table 29: Comparative Sequences relating to SAG1641 (YaeC family protein)

>SEQ ID NO 2954:35_A909 frame: 1
NQEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
KVIKDTSADIPQW

>SEQ ID NO 2955:35_CJB110 frame: 2
SKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANKDVDINAFQHY
NFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDATNGSRALYVL
QSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAIINNTYIEQANL
KPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVKKVIKDTSADI
PQW

>SEQ ID NO 2956:35_COH1 frame: 2
VSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANKDVD
INAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDATNG
SRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAIINNT
YIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVKKVI
KDTSADIPQW

>SEQ ID NO 2957:35_H36B frame: 3
EVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANKDV
DINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDATN
GSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAIINN
TYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVKKV
IKDTSADIPQW

>SEQ ID NO 2958:35_JM9130013 frame: 2
SASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANKDVDI
NAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDATNGS
RALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAIINNTY
IEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVKKVIK
DTSADIPQW

>SEQ ID NO 2959:35_M732 frame: 1
NQEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
KVIKD

>SEQ ID NO 2960:35_M781 frame: 2
VSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANKDVD
INAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDATNG
SRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAIINNT
YIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAIWDAYHTDEVKKVI
KDTSADIPQW

```
SEQ2950      QEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
SEQ2951      QEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
SEQ2952      QEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
SEQ2953      QEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
SEQ2954      QEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
SEQ2955      ---------SKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
SEQ2956      --VSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
SEQ2957      -EVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
SEQ2958      ---SASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
SEQ2959      QEVSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK
SEQ2960      --VSASSTSSKVVKVGVMTFSDTEKARWDKIEKLVGDKAKIKFTEFTDYTQPNQATANK

SEQ2950      DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
SEQ2951      DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
SEQ2952      DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
SEQ2953      DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
SEQ2954      DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
SEQ2955      DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
SEQ2956      DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
SEQ2957      DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
SEQ2958      DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
```

Table 29: Comparative Sequences relating to SAG1641 (YaeC family protein)

```
SEQ2959          DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA
SEQ2960          DVDINAFQHYNFLENWNKENKKNLIPLEKTYLAPIRIYSEKVKSLKKLKKGATIAIPNDA

SEQ2950          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
SEQ2951          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
SEQ2952          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
SEQ2953          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
SEQ2954          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
SEQ2955          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
SEQ2956          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
SEQ2957          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
SEQ2958          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
SEQ2959          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII
SEQ2960          TNGSRALYVLQSAGLIKLNVSGKKVATVANITSNKKDINIQELDASQTPRALKDVDAAII

SEQ2950          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
SEQ2951          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
SEQ2952          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
SEQ2953          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
SEQ2954          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
SEQ2955          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
SEQ2956          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
SEQ2957          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
SEQ2958          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
SEQ2959          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAILDAYHTDEVK
SEQ2960          NNTYIEQANLKPSDAIFVEKSDKNSKQWINIIAGRKNWKKQKNAKAIQAIWDAYHTDEVK

SEQ2950          KVIKDTSADIPQWNPAFLY
SEQ2951          KVIKDTSADIPQW------
SEQ2952          KVIKDTSADIP--------
SEQ2953          KVIKDTSADIPQW------
SEQ2954          KVIKDTSADIPQW------
SEQ2955          KVIKDTSADIPQW------
SEQ2956          KVIKDTSADIPQW------
SEQ2957          KVIKDTSADIPQW------
SEQ2958          KVIKDTSADIPQW------
SEQ2959          KVIKD--------------
SEQ2960          KVIKDTSADIPQW------
```

Table 30: Comparative Sequences relating to SAG2147 (protein of uknown function / lipoprotein, putative)

SEQ ID NO. 3001: SAG2147 FROM THE 1169NT1 GBS TYPE V STRAIN (REVERSE COMPLEMENT)

AAAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGCC
AAAAAATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCT
CCAAAACCTTCTCAGGCATCTAATGAAGTCCCAAAATCAAGTTCTCAATCTACAGAAGCT
AATTCTCAGCAACAAGTTACTGCGAGTGAAGAGGCGGCTGTAGAACAAGCAGTTGTAACA
GAAAATACCCCTGCTACCAGTCAGGCACAACAAACTTATGCTGTTACTGAGACAACTTAC
AAACCTGCTCAACACCAGACAAGTGGCCAAGTATTGAGCAATGGAAATACTGCAGGGGCG
GTCGGATCTGCTGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTTGG
GAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAGCT
TCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGATCAAGTT
AATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGGTTAC

SEQ ID NO. 3002: SAG2147 FROM THE 18RS21 GBS TYPE II STRAIN (REVERSE COMPLEMENT)

AAAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAAAGTTC
GCGTAGCCAAAAAATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAA
AACAGGCTCCAAAACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATCTA
CAGAAGCTAATTCTCAGCAACAAGTTACTGCGAGTGAAGAGGCAGCTGTAGAACAAGCAG
TTGTAACAGAAAACACCCCTGCTACCAGTCAGGCACAACAAGCTTATGCTGTTACTGAGA
CAACTTATAGACCTGCTCAACACCAGACGAGTGGCCAAGTATTGAGTAATGGAAATACTG
CAGGGGCTATTGGCTCAGCAGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGT
CTACTTGGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCT
CAGGAGCTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGG
ATCAAGTTAATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGGTTAC

SEQ ID NO. 3003: SAG2147 FROM THE 2603 V/R GBS TYPE V STRAIN (REVERSE COMPLEMENT)

AAAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAAAGT
TCGCGTAGCCAAAAAATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGT
AAAACAGGCTCCAAAACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATC
TACAGAAGCTAATTCTCAGCAACAAGTTACTGCGAGTGAAGAGGCAGCTGTAGAACAAGC
AGTTGTAACAGAAAACACCCCTGCTACCAGTCAGGCACAACAAGCTTATGCTGTTACTGA
GACAACTTATAGACCTGCTCAACACCAGACGAGTGGCCAAGTATTGAGTAATGGAAATAC
TGCAGGGGCTATTGGCTCAGCAGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCA
GTCTACTTGGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGC
CTCAGGAGCTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCA
GGATCAAGTTAATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGGTTA
C

SEQ ID NO. 3004: SAG2147 FROM THE 090 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)

TAGCCAAAAAATCAAAAATGATTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAAC
AGGCTCCAAAACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAG
AAGCTAATTCTCAGCAACAAGTTACTGCGAGTGAAGAGGCAGCTGTAGAACAAGCAGTTG
TAACAGAAAACACCCCTGCTACCAGTCAGGCACAACAAGCTTATGCTGTTACTGAGACAA
CTTATAGACCTGCTCAACACCAGACGAGTGGCCAAGTATTGAGTAATGGAAATACTGCAG
GGGCTATTGGCTCAGCAGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTA
CTTGGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAG
GAGCTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGA

SEQ ID NO. 3005: SAG2147 FROM THE A909 GBS TYPE Ia STRAIN (REVERSE COMPLEMENT)

AAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCAAAACCTTCTCAGGCA
TCTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAATTCTCAGCAACAAGTT
ACTGCGAGTGAAGAGGCAGCTGTAGAACAAGCAGTTGTAACAGAAAACACCCCTGCTACC
AGTCAGGCACAACAAGCTTATGCTGTTACTGAGACAACTTATAGACCTGCTCAACACCAG
ACAAGTGGCCAAGTATTGAGTAATGGAAATACTGCAGGGGCTATTGGCTCAGCAGCTGCA
GCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTTGGGAACATATTATTGCCCGT
GAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAGCTTCAGGACTTTTCCAAACG
ATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGAATCAAGTTAATTCAGCTATTAAAGCT
TATCGTGCTCAAGGTTTATCA

Table 30: Comparative Sequences relating to SAG2147 (protein of uknown function / lipoprotein, putative)

SEQ ID NO. 3006: SAG2147 FROM THE CJB110 GBS NONTYPEABLE STRAIN (REVERSE COMPLEMENT)

AATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGCCAAAAAATCAAAAATGACTAAGGCGA
CATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCAAAACCTTCTCAGGCATCTAATG
AAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAATTCTCAGCAACAAGTTACTGCGA
GTGAAGAGGCAGCTGTAGAACAAGCAGTTGTAACAGAAAACACCCCTGCTACCAGTCAGG
CACAACAAGCTTATGCTGTTACTGAGACAACTTATAGACCTGCTCAACACCAGACGAGTG
GCCAAGTATTGAGTAATGGAAATACTGCAGGGGCTATTGGCTCAGCAGCTGCAGCACAAA
TGGCTGCTGCAACAGGAGTCCCTCAGTCTACTTGGGAACATATTATTGCCCGTGAATCAA
ATGGTAATCCTAATGTTGCTAATGCCTCAGGAGCTTCAGGACTTTTCCAAACGATGCCAG
GTTGGGGTTCAACAGCTACAGTTCAGGATCAAGTTAATTCAGCTATTAAAGCTTATCGTG
CTCAAGGTTTATCAGCTTGGGGTTAC

SEQ ID NO. 3007: SAG2147 FROM THE COH1 GBS TYPE III STRAIN (REVERSE COMPLEMENT)

AAAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAA
AGTTCGCGTAGCCAAAAAATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGA
TGTAAAACAGGCTCCAAAACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCA
ATCTACAGAAGCTAATTCTCAGCAACAAGTTACTGCGAGTGAAGAGGCGGCTGTAGAACA
AGCAGTTGTAACAGAAAATACCCCTGCTACCAGTCAGGCACAACAAACTTATGCTGTTAC
TGAGACAACTTACAAACCTGCTCAACACCAGACAAGTGGCCAAGTATTGAGCAATGGAAA
TACTGCAGGGGCGGTCGGATCTGCTGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCC
TCAGTCTACTTGGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAA
TGCCTCAGGAGCTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGT
TCAGGATCAAGTTAATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGG
TTAC

SEQ ID NO. 3008: SAG2147 FROM THE H36b GBS TYPE Ib STRAIN (REVERSE COMPLEMENT)

AAAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGC
AGATAAAGTTCGCGTAGCCAAAAAATCAAAAATGACTAAGGCGACATCTAAATCAAAAGT
AGAAGATGTAAAACAGGCTCCAAAACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAG
TTCTCAATCTACAGAAGCTAATTCTCAGCAACAAGTTACTGCGAGTGAAGAGGCAGCTGT
AGAACAAGCAGTTGTAACAGAAAACACCCCTGCTACCAGTCAGGCACAACAAGCTTATGC
TGTTACTGAGACAACTTATAGACCTGCTCAACACCAGACAAGTGGCCAAGTATTGAGTAA
TGGAAATACTGCAGGGGCTATTGGCTCAGCAGCTGCAGCACAAATGGCTGCTGCAACAGG
AGTCCCTCAGTCTACTTGGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGT
TGCTAATGCCTCAGGAGCTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGC
TACAGTTCAGGATCAAGTTAATTCAGCTATTAAAGCTT

SEQ ID NO. 3009: SAG2147 FROM THE M732 GBS TYPE III STRAIN (REVERSE COMPLEMENT)

AAAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGC
CAAAAAATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGC
TCCAAAACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGC
TAATTCTCAGCAACAAGTTACTGCGAGTGAAGAGGCGGCTGTAGAACAAGCAGTTGTAAC
AGAAAATACCCCTGCTACCAGTCAGGCACAACAAACTTATGCTGTTACTGAGACAACTTA
CAAACCTGCTCAACACCAGACAAGTGGCCAAGTATTGAGCAATGGAAATACTGCAGGGGC
GGTCGGATCTGCTGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTTG
GGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAGC
TTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGATCAAGT
TAATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGGTTA

Table 30: Comparative Sequences relating to SAG2147 (protein of uknown function / lipoprotein, putative)

SEQ ID NO. 3010: SAG2147 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)

```
GTAACCCCAAGCTGATAAACCTTGAGCACGATAAGCTTTAATAGCTGAATTAACTTGATC
CTGAACTGTAGCTGTTGAACCCCAACCTGGCATCGTTTGGAAAAGTCCTGAAGCTCCTGA
GGCATTAGCAACATTAGGATTACCATTTGATTCACGGGCAATAATATGTTCCCAAGTAGA
CTGAGGGACTCCTGTTGCAGCAGCCATTTGTGCTGCAGCAGCAGATCCGACCGCCCCTGC
AGTATTTCCATTGCTCAATACTTGGCCACTTGTCTGGTGTTGAGCAGGTTTGTAAGTTGT
CTCAGTAACAGCATAAGTTTGTTGTGCCTGACTGGTAGCAGGGGTATTTTCTGTTACAAC
TGCTTGTTCTACAGCCGCCTCTTCACTCGCAGTAACTTGTTGCTGAGAATTAGCTTCTGT
AGATTGAGAACTTGATTTTGGGGCTTCATTAGATGCCTGAGAAGGTTTTGGAGCCTGTTT
TACATCTTCTACTTTTGATTTAGATGTCGCCTTAGTCATTTTTGATTTTTTGGCTACGCG
AACTTTATCTGCTTTTGACAAAGA
```

```
SEQ3001          ------------------------------------------------------------
SEQ3002          ------------------------------------------------------------
SEQ3003          ------------------------------------------------------------
SEQ3004          ------------------------------------------------------------
SEQ3005          AGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCAAAACCTTCTCAGGCA
SEQ3007          ------------------------------------------------------------
SEQ3008          ------------------------------------------------------------
SEQ3009          ------------------------------------------------------------
SEQ3010          ------------------------------------------------------------

SEQ3001          ------------------------------------------------------------
SEQ3002          ------------------------------------------------------------
SEQ3003          ------------------------------------------------------------
SEQ3004          ------------------------------------------------------------
SEQ3005          CTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAATTCTCAGCAACAAGTT
SEQ3007          ------------------------------------------------------------
SEQ3008          ------------------------------------------------------------
SEQ3009          ------------------------------------------------------------
SEQ3010          ------------------------------------------------------------

SEQ3001          ------------------------------------------------------------
SEQ3002          ------------------------------------------------------------
SEQ3003          ------------------------------------------------------------
SEQ3004          ------------------------------------------------------------
SEQ3005          CTGCGAGTGAAGAGGCAGCTGTAGAACAAGCAGTTGTAACAGAAAACACCCCTGCTACC
SEQ3007          ------------------------------------------------------------
SEQ3008          ------------------------------------------------------------
SEQ3009          ------------------------------------------------------------
SEQ3010          ------------------------------------------------------------

SEQ3001          ------------------------------------------------------------
SEQ3002          ------------------------------------------------------------
SEQ3003          ------------------------------------------------------------
SEQ3004          ------------------------------------------------------------
SEQ3005          GTCAGGCACAACAAGCTTATGCTGTTACTGAGACAACTTATAGACCTGCTCAACACCAG
SEQ3007          ------------------------------------------------------------
SEQ3008          ------------------------------------------------------------
SEQ3009          ------------------------------------------------------------
SEQ3010          ------------------------------------------------------------

SEQ3001          ------------------------------------------------------------
SEQ3002          ------------------------------------------------------------
SEQ3003          ------------------------------------------------------------
SEQ3004          ------------------------------------------------------------
SEQ3005          CAAGTGGCCAAGTATTGAGTAATGGAAATACTGCAGGGGCTATTGGCTCAGCAGCTGCA
SEQ3007          ------------------------------------------------------------
SEQ3008          ------------------------------------------------------------
SEQ3009          ------------------------------------------------------------
SEQ3010          ------------------------------------------------------------
```

Table 30: Comparative Sequences relating to SAG2147 (protein of uknown function / lipoprotein, putative)

```
SEQ3001          ------------------------------------------------------------
SEQ3002          ------------------------------------------------------------
SEQ3003          ------------------------------------------------------------
SEQ3004          ------------------------------------------------------------
SEQ3005          CACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTTGGGAACATATTATTGCCCGT
SEQ3007          ------------------------------------------------------------
SEQ3008          ------------------------------------------------------------
SEQ3009          ------------------------------------------------------------
SEQ3010          ------------------------------------------------------------

SEQ3001          ------------------------------------------------------------
SEQ3002          ------------------------------------------------------------
SEQ3003          ------------------------------------------------------------
SEQ3004          ------------------------------------------------------------
SEQ3005          AATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAGCTTCAGGACTTTTCCAAACG
SEQ3007          ------------------------------------------------------------
SEQ3008          ------------------------------------------------------------
SEQ3009          ------------------------------------------------------------
SEQ3010          ------------------------------------------------------------

SEQ3001          ------------------------------------------------------------
SEQ3002          ------------------------------------------------------------
SEQ3003          ------------------------------------------------------------
SEQ3004          ------------------------------------------------------------
SEQ3005          TGCCAGGTTGGGGTTCAACAGCTACAGTTCAGAATCAAGTTAATTCAGCTATTAAAGCT
SEQ3007          ------------------------------------------------------------
SEQ3008          ------------------------------------------------------------
SEQ3009          ------------------------------------------------------------
SEQ3010          ------------------------------------------------------------

SEQ3001          AAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGCCAAA
SEQ3002          AAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGCCAAA
SEQ3003          AAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGCCAAA
SEQ3004          ----------------------------------------------------TAGCCAAA
SEQ3005          ATCGTGCTCAAGGTTTATCASAATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGCCAAA
SEQ3007          AAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGCCAAA
SEQ3008          AAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGCCAAA
SEQ3009          AAAGTTCACAAGTTACTACTGAATCTTTGTCAAAAGCAGATAAAGTTCGCGTAGCCAAA
SEQ3010          -------GTAACCCCAAGCTGA---TAAACCTTGAGCACGATAAGCTTTAATAGCTGAA

SEQ3001          AATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCA
SEQ3002          AATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCA
SEQ3003          AATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCA
SEQ3004          AATCAAAAATGATTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCA
SEQ3005          AATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCA
SEQ3007          AATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCA
SEQ3008          AATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCA
SEQ3009          AATCAAAAATGACTAAGGCGACATCTAAATCAAAAGTAGAAGATGTAAAACAGGCTCCA
SEQ3010          TAACTTGATCCTGAACTGTAGCTGTTGAACCCCAACCTGGCATCGTTTGGAAAAGTCCT

SEQ3001          AACCTTCTCAGGCATCTAATGAAGTCCCAAAATCAAGTTCTCAATCTACAGAAGCTAAT
SEQ3002          AACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAAT
SEQ3003          AACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAAT
SEQ3004          AACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAAT
SEQ3005          AACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAAT
SEQ3007          AACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAAT
SEQ3008          AACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAAT
SEQ3009          AACCTTCTCAGGCATCTAATGAAGCCCCAAAATCAAGTTCTCAATCTACAGAAGCTAAT
SEQ3010          AAGCTCCTGAGGCATT---AGCAACATTAGGATTAC-CATTTGATTCACGGGCAATAAT

SEQ3001          TCTCAGCAACAAGTTACTGCGAGTGAAGAGGCGGCTGTAGAACAAGCAGTTGTAACAGA
SEQ3002          TCTCAGCAACAAGTTACTGCGAGTGAAGAGGCAGCTGTAGAACAAGCAGTTGTAACAGA
SEQ3003          TCTCAGCAACAAGTTACTGCGAGTGAAGAGGCAGCTGTAGAACAAGCAGTTGTAACAGA
SEQ3004          TCTCAGCAACAAGTTACTGCGAGTGAAGAGGCAGCTGTAGAACAAGCAGTTGTAACAGA
SEQ3005          TCTCAGCAACAAGTTACTGCGAGTGAAGAGGCAGCTGTAGAACAAGCAGTTGTAACAGA
SEQ3007          TCTCAGCAACAAGTTACTGCGAGTGAAGAGGCGGCTGTAGAACAAGCAGTTGTAACAGA
SEQ3008          TCTCAGCAACAAGTTACTGCGAGTGAAGAGGCAGCTGTAGAACAAGCAGTTGTAACAGA
SEQ3009          TCTCAGCAACAAGTTACTGCGAGTGAAGAGGCGGCTGTAGAACAAGCAGTTGTAACAGA
SEQ3010          TGTTCCCAAGTAGACTGAGGGACTCCTGTTGCAGCAGCCATTTGTGCTGCAGCAGCAGA
```

Table 30: Comparative Sequences relating to SAG2147 (protein of uknown function / lipoprotein, putative)

```
SEQ3001          --AAATACCCCTGCTACCAGTCAGGCACAACAAACTTATGCTGTTACTGAGACAACTTA
SEQ3002          --AAACACCCCTGCTACCAGTCAGGCACAACAAGCTTATGCTGTTACTGAGACAACTTA
SEQ3003          --AAACACCCCTGCTACCAGTCAGGCACAACAAGCTTATGCTGTTACTGAGACAACTTA
SEQ3004          --AAACACCCCTGCTACCAGTCAGGCACAACAAGCTTATGCTGTTACTGAGACAACTTA
SEQ3005          --AAACACCCCTGCTACCAGTCAGGCACAACAAGCTTATGCTGTTACTGAGACAACTTA
SEQ3007          --AAATACCCCTGCTACCAGTCAGGCACAACAAACTTATGCTGTTACTGAGACAACTTA
SEQ3008          --AAACACCCCTGCTACCAGTCAGGCACAACAAGCTTATGCTGTTACTGAGACAACTTA
SEQ3009          --AAATACCCCTGCTACCAGTCAGGCACAACAAACTTATGCTGTTACTGAGACAACTTA
SEQ3010          CCGACCGCCCCTGCAGTATTTCCATTGCTCAATACTTG-GCCACTTGTCTGGTGTTGAG

SEQ3001          AAACCTGCTCAACACCAGACAAGTGGC-CAAGTATTGAGCAATGGAAATACTGCAGGGG
SEQ3002          AGACCTGCTCAACACCAGACGAGTGGC-CAAGTATTGAGTAATGGAAATACTGCAGGGG
SEQ3003          AGACCTGCTCAACACCAGACGAGTGGC-CAAGTATTGAGTAATGGAAATACTGCAGGGG
SEQ3004          AGACCTGCTCAACACCAGACGAGTGGC-CAAGTATTGAGTAATGGAAATACTGCAGGGG
SEQ3005          AGACCTGCTCAACACCAGACGAGTGGC-CAAGTATTGAGTAATGGAAATACTGCAGGGG
SEQ3007          AAACCTGCTCAACACCAGACAAGTGGC-CAAGTATTGAGCAATGGAAATACTGCAGGGG
SEQ3008          AGACCTGCTCAACACCAGACAAGTGGC-CAAGTATTGAGTAATGGAAATACTGCAGGGG
SEQ3009          AAACCTGCTCAACACCAGACAAGTGGC-CAAGTATTGAGCAATGGAAATACTGCAGGGG
SEQ3010          AGGTTTGTAAGTTGTCTCAGTAACAGCATAAGTTTGTTGTGCCTGACTGGTAGCAGGGG

SEQ3001          GGTCGGATCTGCTGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTT
SEQ3002          TATTGGCTCAGCAGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTT
SEQ3003          TATTGGCTCAGCAGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTT
SEQ3004          TATTGGCTCAGCAGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTT
SEQ3005          TATTGGCTCAGCAGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTT
SEQ3007          GGTCGGATCTGCTGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTT
SEQ3008          TATTGGCTCAGCAGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTT
SEQ3009          GGTCGGATCTGCTGCTGCAGCACAAATGGCTGCTGCAACAGGAGTCCCTCAGTCTACTT
SEQ3010          A-TTT--TCTGTTACAACTGCTTGTTCTACAGCCGCCTCTTCACTCGCAGTAACTTGTT

SEQ3001          GGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAG
SEQ3002          GGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAG
SEQ3003          GGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAG
SEQ3004          GGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAG
SEQ3005          GGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAG
SEQ3007          GGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAG
SEQ3008          GGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAG
SEQ3009          GGGAACATATTATTGCCCGTGAATCAAATGGTAATCCTAATGTTGCTAATGCCTCAGGAG
SEQ3010          GCTGAGA-ATTAGCTTCTGTAGATTGAG---AA--CTTGATTTTGGGGCTTCATTAGATG

SEQ3001          CTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGATCAAG
SEQ3002          CTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGATCAAG
SEQ3003          CTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGATCAAG
SEQ3004          CTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGA-----
SEQ3005          CTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGATCAAG
SEQ3007          CTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGATCAAG
SEQ3008          CTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGATCAAG
SEQ3009          CTTCAGGACTTTTCCAAACGATGCCAGGTTGGGGTTCAACAGCTACAGTTCAGGATCAAG
SEQ3010          CCTGAGAAGGTTTT-----GGAGCCTGTTTTACATCTTCTACTTTTGATTTAGATGTCGC

SEQ3001          TAATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGGTTAC--
SEQ3002          TAATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGGTTAC--
SEQ3003          TAATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGGTTAC--
SEQ3004          ------------------------------------------------------
SEQ3005          TAATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGGTTAC--
SEQ3007          TAATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGGTTAC--
SEQ3008          TAATTCAGCTATTAAAGCTT----------------------------------
SEQ3009          TAATTCAGCTATTAAAGCTTATCGTGCTCAAGGTTTATCAGCTTGGGGTTA---
SEQ3010          TTAGTCA-TTTTTGATTTTTTGGCTACGCGAACTTTATCTGCTTTTGACAAAGA
```

Table 30: Comparative Sequences relating to SAG2147 (protein of uknown function / lipoprotein, putative)

>SEQ ID NO 3050: 25_1169NT frame: 1
KSSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEVPKSSSQSTEAN
SQQQVTASEEAAVEQAVVTENTPATSQAQQTYAVTETTYKPAQHQTSGQVLSNGNTAGAV
GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SAIKAYRAQGLSAWGY

>SEQ ID NO 3051:25_18RS21 frame: 1
KSSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SQQQVTASEEAAVEQAVVTENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAI
GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SAIKAYRAQGLSAWGY

>SEQ ID NO 3052:25_2603 frame: 1
KSSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SQQQVTASEEAAVEQAVVTENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAI
GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SAIKAYRAQGLSAWGY

>SEQ ID NO 3053:25_090 frame: 3
AKKSKMIKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEANSQQQVTASEEAAVEQAVV
TENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAIGSAAAAQMAAATGVPQST
WEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQ

>SEQ ID NO 3054:25_A909 frame: 1
KATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEANSQQQVTASEEAAVEQAVVTENTPAT
SQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAIGSAAAAQMAAATGVPQSTWEHIIAR
ESNGNPNVANASGASGLFQTMPGWGSTATVQNQVNSAIKAYRAQGLS

>SEQ ID NO 3055:25_CJB110 frame: 3
SLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEANSQQQVTAS
EEAAVEQAVVTENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAIGSAAAAQM
AAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVNSAIKAYRA
QGLSAWGY

>SEQ ID NO 3056:25_COH1 frame: 1
KSSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SQQQVTASEEAAVEQAVVTENTPATSQAQQTYAVTETTYKPAQHQTSGQVLSNGNTAGAV
GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SAIKAYRAQGLSAWGY

>SEQ ID NO 3057:25_H36B frame: 1
KSSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SQQQVTASEEAAVEQAVVTENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAI
GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SAIKA

>SEQ ID NO 3058:25_M732 frame: 1
KSSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SQQQVTASEEAAVEQAVVTENTPATSQAQQTYAVTETTYKPAQHQTSGQVLSNGNTAGAV
GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SAIKAYRAQGLSAWG

>SEQ ID NO 3059:25_M781 frame: 4
SLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEANSQQQVTAS
EEAAVEQAVVTENTPATSQAQQTYAVTETTYKPAQHQTSGQVLSNGNTAGAVGSAAAAQM
AAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVNSAIKAYRA
QGLSAWGY

```
SEQ3050         SSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEVPKSSSQSTEAN
SEQ3051         SSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SEQ3052         SSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SEQ3053         -----------------AKKSKMIKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SEQ3054         ------------------------KATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SEQ3055         -------SLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SEQ3056         SSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SEQ3057         SSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SEQ3058         SSQVTTESLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
SEQ3059         -------SLSKADKVRVAKKSKMTKATSKSKVEDVKQAPKPSQASNEAPKSSSQSTEAN
```

Table 30: Comparative Sequences relating to SAG2147 (protein of uknown function / lipoprotein, putative)

```
SEQ3050         SQQQVTASEEAAVEQAVVTENTPATSQAQQTYAVTETTYKPAQHQTSGQVLSNGNTAGAV
SEQ3051         SQQQVTASEEAAVEQAVVTENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAI
SEQ3052         SQQQVTASEEAAVEQAVVTENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAI
SEQ3053         SQQQVTASEEAAVEQAVVTENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAI
SEQ3054         SQQQVTASEEAAVEQAVVTENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAI
SEQ3055         SQQQVTASEEAAVEQAVVTENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAI
SEQ3056         SQQQVTASEEAAVEQAVVTENTPATSQAQQTYAVTETTYKPAQHQTSGQVLSNGNTAGAV
SEQ3057         SQQQVTASEEAAVEQAVVTENTPATSQAQQAYAVTETTYRPAQHQTSGQVLSNGNTAGAI
SEQ3058         SQQQVTASEEAAVEQAVVTENTPATSQAQQTYAVTETTYKPAQHQTSGQVLSNGNTAGAV
SEQ3059         SQQQVTASEEAAVEQAVVTENTPATSQAQQTYAVTETTYKPAQHQTSGQVLSNGNTAGAV

SEQ3050         GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SEQ3051         GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SEQ3052         GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SEQ3053         GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQ----
SEQ3054         GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQNQVN
SEQ3055         GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SEQ3056         GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SEQ3057         GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SEQ3058         GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN
SEQ3059         GSAAAAQMAAATGVPQSTWEHIIARESNGNPNVANASGASGLFQTMPGWGSTATVQDQVN

SEQ3050         AIKAYRAQGLSAWGY
SEQ3051         AIKAYRAQGLSAWGY
SEQ3052         AIKAYRAQGLSAWGY
SEQ3053         ---------------
SEQ3054         AIKAYRAQGLS----
SEQ3055         AIKAYRAQGLSAWGY
SEQ3056         AIKAYRAQGLSAWGY
SEQ3057         AIKA-----------
SEQ3058         AIKAYRAQGLSAWG-
SEQ3059         AIKAYRAQGLSAWGY
```

Table 31: Comparative Sequences relating to SAG2148 (LysM domain protein)

SEQ ID NO. 3101: SAG2148 FROM THE 1169NT1 GBS TYPE V STRAIN

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAAGAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTACAGCTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCTGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGCTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

SEQ ID NO. 3102: SAG2148 FROM THE 18RS21 GBS TYPE II STRAIN

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAAGAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTACAGCTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGTTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

SEQ ID NO. 3103: SAG2148 FROM THE 2603 V/R GBS TYPE V STRAIN

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAAGAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTACAGCTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGTTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

SEQ ID NO. 3104: SAG2148 FROM THE 090 GBS TYPE Ia STRAIN

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAAGAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTAAAGCTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGTTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

SEQ ID NO. 3105: SAG2148 FROM THE A909 GBS TYPE Ia STRAIN

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAAGAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTACAGCTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGCTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

SEQ ID NO. 3106: SAG2148 FROM THE CJB110 GBS NONTYPEABLE STRAIN

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAAGAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTAAAGCTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGTTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

SEQ ID NO. 3107: SAG2148 FROM THE COH1 GBS TYPE III STRAIN

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAATAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTACAGCTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCTGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGCTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

Table 31: Comparative Sequences relating to SAG2148 (LysM domain protein)

SEQ ID NO. 3108: SAG2148 FROM THE H36b GBS TYPE Ib STRAIN (REVERSE COMPLEMENT)

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAAGAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTACAGCTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGCTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

SEQ ID NO. 3109: SAG2148 FROM THE JM9130013 GBS TYPE VIII STRAIN (REVERSE COMPLEMENT)

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAAGAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGACGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTACAACTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGCTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

SEQ ID NO. 3110: SAG2148 FROM THE M732 GBS TYPE III STRAIN

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAATAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTACAGCTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCTGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGCTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

SEQ ID NO. 3111: SAG2148 FROM THE M781 GBS TYPE III STRAIN (REVERSE COMPLEMENT)

GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACTACGGTACAATAGTTAGTG
TCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGTGATGTTTTAAAATTGGATAATTCTACAGCTAGTCAA
GCAGAAGCAAAATCTCAACCAACAATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCTGCA
AAAGAAGAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGAAGATATCAACTG
TCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAAGTAGCGGACAATTATGTGGCTTCTCGTTAC
GGATCTTGGTCGGCAGCGCTATCATTTTGGAATAGTAACGGCTGGTAT

SEQ3101	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT
SEQ3102	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT
SEQ3103	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT
SEQ3104	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT
SEQ3105	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT
SEQ3106	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT
SEQ3107	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT
SEQ3108	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT
SEQ3109	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT
SEQ3110	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT
SEQ3111	GCATCTTATACCGTGAAATCAGGTGATACCTTATCAGCTATTGCTAAAAATCATAAAACT

SEQ3101	ACGGTACAAGAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGT
SEQ3102	ACGGTACAAGAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGT
SEQ3103	ACGGTACAAGAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGT
SEQ3104	ACGGTACAAGAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGT
SEQ3105	ACGGTACAAGAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGT
SEQ3106	ACGGTACAAGAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGT
SEQ3107	ACGGTACAATAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGT
SEQ3108	ACGGTACAAGAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGT
SEQ3109	ACGGTACAAGAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGACGTCATCAGTATAGGT
SEQ3110	ACGGTACAATAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGT
SEQ3111	ACGGTACAATAGTTAGTGTCTCTCAATAGTATCAGTAACGCTGATGTCATCAGTATAGGT

Table 31: Comparative Sequences relating to SAG2148 (LysM domain protein)

```
SEQ3101         GATGTTTTAAAATTGGATAATTCTACAGCTAGTCAAGCAGAAGCAAAATCTCAACCAACA
SEQ3102         GATGTTTTAAAATTGGATAATTCTACAGCTAGTCAAGCAGAAGCAAAATCTCAACCAACA
SEQ3103         GATGTTTTAAAATTGGATAATTCTACAGCTAGTCAAGCAGAAGCAAAATCTCAACCAACA
SEQ3104         GATGTTTTAAAATTGGATAATTCTAAAGCTAGTCAAGCAGAAGCAAAATCTCAACCAACA
SEQ3105         GATGTTTTAAAATTGGATAATTCTACAGCTAGTCAAGCAGAAGCAAAATCTCAACCAACA
SEQ3106         GATGTTTTAAAATTGGATAATTCTAAAGCTAGTCAAGCAGAAGCAAAATCTCAACCAACA
SEQ3107         GATGTTTTAAAATTGGATAATTCTACAGCTAGTCAAGCAGAAGCAAAATCTCAACCAACA
SEQ3108         GATGTTTTAAAATTGGATAATTCTACAGCTAGTCAAGCAGAAGCAAAATCTCAACCAACA
SEQ3109         GATGTTTTAAAATTGGATAATTCTACAACTAGTCAAGCAGAAGCAAAATCTCAACCAACA
SEQ3110         GATGTTTTAAAATTGGATAATTCTACAGCTAGTCAAGCAGAAGCAAAATCTCAACCAACA
SEQ3111         GATGTTTTAAAATTGGATAATTCTACAGCTAGTCAAGCAGAAGCAAAATCTCAACCAACA

SEQ3101         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCTGCAAAAGAA
SEQ3102         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCAAAAGAA
SEQ3103         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCAAAAGAA
SEQ3104         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCAAAAGAA
SEQ3105         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCAAAAGAA
SEQ3106         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCAAAAGAA
SEQ3107         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCTGCAAAAGAA
SEQ3108         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCAAAAGAA
SEQ3109         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCCGCAAAAGAA
SEQ3110         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCTGCAAAAGAA
SEQ3111         ATTGAAAATTCAATGAATTCTTCATCAAATTTGAGTTCAAGTGATTCAGCTGCAAAAGAA

SEQ3101         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA
SEQ3102         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA
SEQ3103         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA
SEQ3104         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA
SEQ3105         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA
SEQ3106         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA
SEQ3107         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA
SEQ3108         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA
SEQ3109         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA
SEQ3110         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA
SEQ3111         GAAATAGCTCGTCGTGAATCAAATGGTAGTTATACTGCACAGAATGGACAATATTATGGA

SEQ3101         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA
SEQ3102         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA
SEQ3103         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA
SEQ3104         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA
SEQ3105         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA
SEQ3106         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA
SEQ3107         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA
SEQ3108         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA
SEQ3109         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA
SEQ3110         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA
SEQ3111         AGATATCAACTGTCTCAATCTTACCTAAATGGCGACTTATCTCCTGAAAATCAAGAAAAA

SEQ3101         GTAGCGGACAATTATGTGGCTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
SEQ3102         GTAGCGGACAATTATGTGGTTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
SEQ3103         GTAGCGGACAATTATGTGGTTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
SEQ3104         GTAGCGGACAATTATGTGGTTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
SEQ3105         GTAGCGGACAATTATGTGGCTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
SEQ3106         GTAGCGGACAATTATGTGGTTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
SEQ3107         GTAGCGGACAATTATGTGGCTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
SEQ3108         GTAGCGGACAATTATGTGGCTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
SEQ3109         GTAGCGGACAATTATGTGGCTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
SEQ3110         GTAGCGGACAATTATGTGGCTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
SEQ3111         GTAGCGGACAATTATGTGGCTTCTCGTTACGGATCTTGGTCGGCAGCGCTATCATTTTGG
```

Table 31: Comparative Sequences relating to SAG2148 (LysM domain protein)

SEQ3101	AATAGTAACGGCTGGTAT
SEQ3102	AATAGTAACGGCTGGTAT
SEQ3103	AATAGTAACGGCTGGTAT
SEQ3104	AATAGTAACGGCTGGTAT
SEQ3105	AATAGTAACGGCTGGTAT
SEQ3106	AATAGTAACGGCTGGTAT
SEQ3107	AATAGTAACGGCTGGTAT
SEQ3108	AATAGTAACGGCTGGTAT
SEQ3109	AATAGTAACGGCTGGTAT
SEQ3110	AATAGTAACGGCTGGTAT
SEQ3111	AATAGTAACGGCTGGTAT

>SEQ ID NO 3150:15_1169NT frame: 1
ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVASRYGSWSAALSFWNSNGWY

>SEQ ID NO 3151:15_18RS21 frame: 1
ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVVSRYGSWSAALSFWNSNGWY

>SEQ ID NO 3152:15_2603 frame: 1
ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVVSRYGSWSAALSFWNSNGWY

>SEQ ID NO 3153:15_090 frame: 1
ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSKASQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVVSRYGSWSAALSFWNSNGWY

>SEQ ID NO 3154:15_A909 frame: 1
ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVASRYGSWSAALSFWNSNGWY

>SEQ ID NO 3155:15_CJB110 frame: 1
ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSKASQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVVSRYGSWSAALSFWNSNGWY

>SEQ ID NO 3156:15_COH1 frame: 1
ASYTVKSGDTLSAIAKNHKTTVQ.LVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVASRYGSWSAALSFWNSNGWY

>SEQ ID NO 3157:15_H36B frame: 1
ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVASRYGSWSAALSFWNSNGWY

>SEQ ID NO 3158:15_JM9130013 frame: 1
ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTTSQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVASRYGSWSAALSFWNSNGWY

>SEQ ID NO 3159:15_M732 frame: 1
ASYTVKSGDTLSAIAKNHKTTVQ.LVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVASRYGSWSAALSFWNSNGWY

>SEQ ID NO 3160:15_M781 frame: 1
ASYTVKSGDTLSAIAKNHKTTVQ.LVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
VADNYVASRYGSWSAALSFWNSNGWY

Table 31: Comparative Sequences relating to SAG2148 (LysM domain protein)

```
SEQ3150          ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
SEQ3151          ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
SEQ3152          ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
SEQ3153          ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSKASQAEAKSQPT
SEQ3154          ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
SEQ3155          ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSKASQAEAKSQPT
SEQ3156          ASYTVKSGDTLSAIAKNHKTTVQ-LVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
SEQ3157          ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
SEQ3158          ASYTVKSGDTLSAIAKNHKTTVQELVSLNSISNADVISIGDVLKLDNSTTSQAEAKSQPT
SEQ3159          ASYTVKSGDTLSAIAKNHKTTVQ-LVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT
SEQ3160          ASYTVKSGDTLSAIAKNHKTTVQ-LVSLNSISNADVISIGDVLKLDNSTASQAEAKSQPT

SEQ3150          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
SEQ3151          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
SEQ3152          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
SEQ3153          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
SEQ3154          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
SEQ3155          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
SEQ3156          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
SEQ3157          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
SEQ3158          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
SEQ3159          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK
SEQ3160          IENSMNSSSNLSSSDSAAKEEIARRESNGSYTAQNGQYYGRYQLSQSYLNGDLSPENQEK

SEQ3150          VADNYVASRYGSWSAALSFWNSNGWY
SEQ3151          VADNYVVSRYGSWSAALSFWNSNGWY
SEQ3152          VADNYVVSRYGSWSAALSFWNSNGWY
SEQ3153          VADNYVVSRYGSWSAALSFWNSNGWY
SEQ3154          VADNYVASRYGSWSAALSFWNSNGWY
SEQ3155          VADNYVVSRYGSWSAALSFWNSNGWY
SEQ3156          VADNYVASRYGSWSAALSFWNSNGWY
SEQ3157          VADNYVASRYGSWSAALSFWNSNGWY
SEQ3158          VADNYVASRYGSWSAALSFWNSNGWY
SEQ3159          VADNYVASRYGSWSAALSFWNSNGWY
SEQ3160          VADNYVASRYGSWSAALSFWNSNGWY
```

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00003	SAG0017	447	PcsB protein
ORF00004	SAG0018	322	ribose-phosphate pyrophosphokinase
ORF00005	SAG0019	391	aminotransferase, class I
ORF00006	SAG0020	253	recombination protein O
ORF00008	SAG0021	283	protease, putative
ORF00009	SAG0022	330	fatty acid/phospholipid synthesis protein PlsX
ORF00010	SAG0023	79	acyl carrier protein
ORF00011	SAG0024	234	phosphoribosylaminoimidazole-succinocarboxamide synthase
ORF00012	SAG0025	1241	phosphoribosylformylglycinamidine synthase, putative
ORF00013	SAG0026	484	amidophosphoribosyltransferase
ORF00014	SAG0027	340	phosphoribosylformylglycinamidine cyclo-ligase
ORF00015	SAG0028	182	phosphoribosylglycinamide formyltransferase
ORF00016	SAG0029	250	acetyltransferase, GNAT family
ORF00017	SAG0030	515	phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase
ORF00018	SAG0031	283	peptidase, M23/M37 family
ORF00020	SAG0032	434	group B streptococcal surface immunogenic protein
ORF00021	SAG0033	232	N-acetylmannosamine-6-P epimerase, putative
ORF00022	SAG0034	438	sugar ABC transporter, sugar-binding protein
ORF00023	SAG0035	295	sugar ABC transporter, permease protein
ORF00024	SAG0036	276	sugar ABC transporter, permease protein
ORF00025	SAG0037	147	conserved hypothetical protein
ORF00026	SAG0038	220	conserved hypothetical protein
ORF00027	SAG0039	305	N-acetylneuraminate lyase, putative
ORF00028	SAG0040	293	ROK family protein
ORF00029	SAG0041	325	acetyl xylan esterase, putative
ORF00030	SAG0042	267	phosphosugar-binding transcriptional regulator, RpiR family, putative
ORF00031	SAG0043	421	phosphoribosylamine--glycine ligase
ORF00032	SAG0044	162	phosphoribosylaminoimidazole carboxylase, catalytic subunit
ORF00033	SAG0045	363	phosphoribosylaminoimidazole carboxylase, ATPase subunit
ORF00035	SAG0046	463	hypothetical protein
ORF00036	SAG0047	432	adenylosuccinate lyase
ORF00037	SAG0048	303	transcriptional regulator, Cro/CI family
ORF00038	SAG0049	332	Holliday junction DNA helicase RuvB
ORF00039	SAG0050	145	phosphotyrosine protein phosphatase, low molecular weight
ORF00040	SAG0051	126	MORN motif family protein
ORF00041	SAG0052	592	membrane protein, putative
ORF00042	SAG0053	880	aldehyde-alcohol dehydrogenase
ORF00043	SAG0054	338	alcohol dehydrogenase, propanol-preferring
ORF00044	SAG0055	496	threonine synthase
ORF00045	SAG0056	412	MATE efflux family protein
ORF00046	SAG0057	102	ribosomal protein S10
ORF00047	SAG0058	208	ribosomal protein L3
ORF00048	SAG0059	207	ribosomal protein L4
ORF00049	SAG0060	98	ribosomal protein L23
ORF00050	SAG0061	277	ribosomal protein L2
ORF00052	SAG0062	92	ribosomal protein S19
ORF00054	SAG0063	114	ribosomal protein L22
ORF00055	SAG0064	217	ribosomal protein S3

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00056	SAG0065	137	ribosomal protein L16
ORF00058	SAG0066	68	ribosomal protein L29
ORF00059	SAG0067	86	ribosomal protein S17
ORF00060	SAG0068	122	ribosomal protein L14
ORF00061	SAG0069	101	ribosomal protein L24
ORF00063	SAG0070	180	ribosomal protein L5
ORF00064	SAG0071	61	ribosomal protein S14, putative
ORF00065	SAG0072	132	ribosomal protein S8
ORF00066	SAG0073	178	ribosomal protein L6
ORF00068	SAG0074	118	ribosomal protein L18
ORF00069	SAG0075	164	ribosomal protein S5
ORF00070	SAG0076	59	ribosomal protein L30
ORF00071	SAG0077	146	ribosomal protein L15
ORF00072	SAG0078	434	preprotein translocase, SecY subunit
ORF00073	SAG0079	212	adenylate kinase
ORF00074	SAG0080	72	translation initiation factor IF-1
ORF00075	SAG0081	38	ribosomal protein L36
ORF00077	SAG0082	121	ribosomal protein S13
ORF00078	SAG0083	118	ribosomal protein S11
ORF00080	SAG0084	312	DNA-directed RNA polymerase, alpha subunit
ORF00081	SAG0085	128	ribosomal protein L17
ORF00087	SAG0086	97	hypothetical protein
ORF00088	SAG0087	59	hypothetical protein
ORF00089	SAG0088	56	hypothetical protein
ORF00090	SAG0089	183	conserved hypothetical protein
ORF00091	SAG0090	139	conserved hypothetical protein
ORF00093	SAG0091	144	transcriptional regulator ComX1, putative
ORF00094	SAG0092	230	phosphoglycerate mutase family protein
ORF00095	SAG0093	250	D-alanyl-D-alanine carboxypeptidase family protein
ORF00096	SAG0094	191	N-acetylmuramoyl-L-alanine amidase, family 4 protein
ORF00097	SAG0095	344	heat-inducible transcription repressor HrcA
ORF00098	SAG0096	190	heat shock protein GrpE
ORF00099	SAG0097	609	dnaK protein
ORF00100	SAG0098	379	dnaJ protein
ORF00101	SAG0099	415	transcriptional regulator, GntR family
ORF00102	SAG0100	258	tRNA pseudouridine synthase A
ORF00103	SAG0101	252	phosphomethylpyrimidine kinase, putative
ORF00104	SAG0102	154	conserved hypothetical protein
ORF00105	SAG0103	189	conserved hypothetical protein
ORF00106	SAG0104	280	conserved hypothetical protein
ORF00107	SAG0105	427	trigger factor
ORF00108	SAG0106	191	DNA-directed RNA polymerase, delta subunit, putative
ORF00109	SAG0107	534	CTP synthase
ORF00110	SAG0108	308	conserved hypothetical protein
ORF00111	SAG0109	148	deoxyuridine 5`-triphosphate nucleotidohydrolase
ORF00112	SAG0110	454	DNA repair protein RadA
ORF00113	SAG0111	165	carbonic anhydrase-related protein
ORF00115	SAG0112	439	pyridine nucleotide-disulphide oxidoreductase family protein
ORF00116	SAG0113	484	glutamyl-tRNA synthetase
ORF00117	SAG0114	322	ribose ABC transporter, periplasmic D-ribose-binding protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00118	SAG0115	310	ribose ABC transporter, permease protein
ORF00119	SAG0116	492	ribose ABC transporter, ATP-binding protein
ORF00120	SAG0117	132	ribose ABC transporter protein RbsD
ORF00121	SAG0118	303	ribokinase
ORF00122	SAG0119	328	ribose operon repressor RbsR
ORF00123	SAG0120	32	hypothetical protein
ORF00124	SAG0121	362	permease, putative
ORF00125	SAG0122	228	ABC transporter, ATP-binding protein
ORF00126	SAG0123	223	DNA-binding response regulator
ORF00128	SAG0124	356	sensor histidine kinase
ORF00129	SAG0125	396	argininosuccinate synthase
ORF00130	SAG0126	462	argininosuccinate lyase
ORF00131	SAG0127	293	fructose-bisphosphate aldolase
ORF00132	SAG0128	305	L-2-hydroxyisocaproate dehydrogenase
ORF00133	SAG0129	62	ribosomal protein L28
ORF00134	SAG0130	121	conserved hypothetical protein
ORF00135	SAG0131	543	DAK2 domain protein
ORF00136	SAG0132	294	SPFH domain/Band 7 family protein
ORF00137	SAG0133	38	conserved hypothetical protein
ORF00138	SAG0134	96	hypothetical protein
ORF00141	SAG0135	246	amino acid ABC transporter, ATP-binding protein
ORF00142	SAG0136	516	amino acid ABC transporter, amino acid-binding protein/permease protein
ORF00143	SAG0137	627	conserved hypothetical protein
ORF00145	SAG0138	279	undecaprenol kinase, putative
ORF00146	SAG0139	251	negative regulator of competence MecA, putative
ORF00148	SAG0140	386	glycosyl transferase, group 4 family protein
ORF00149	SAG0141	256	ABC transporter, ATP-binding protein
ORF00150	SAG0142	420	conserved hypothetical protein
ORF00151	SAG0143	410	selenocysteine lyase
ORF00152	SAG0144	147	NifU family protein
ORF00153	SAG0145	472	conserved hypothetical protein
ORF00154	SAG0146	395	penicillin-binding protein 4, putative
ORF00155	SAG0147	411	D-alanyl-D-alanine carboxypeptidase
ORF00156	SAG0148	551	oligopeptide ABC transporter, substrate binding protein, putative
ORF00157	SAG0149	304	oligopeptide ABC transporter, permease protein
ORF00158	SAG0150	343	oligopeptide ABC transporter, permease protein
ORF00160	SAG0151	348	oligopeptide ABC transporter, ATP-binding protein
ORF00161	SAG0152	310	oligopeptide ABC transporter, ATP-binding protein
ORF00166	SAG0153	283	4-diphosphocytidyl-2C-methyl-D-erythritol kinase
ORF00167	SAG0154	147	adc operon repressor AdcR
ORF00168	SAG0155	236	zinc ABC transporter, ATP-binding protein
ORF00169	SAG0156	270	zinc ABC transporter, permease protein
ORF00172	SAG0158	419	tyrosyl-tRNA synthetase
ORF00173	SAG0159	765	penicillin-binding protein 1B, putative
ORF00174	SAG0160	1191	DNA-directed RNA polymerase, beta subunit
ORF00176	SAG0161	1216	DNA-directed RNA polymerase beta' subunit
ORF00178	SAG0162	121	conserved hypothetical protein
ORF00179	SAG0163	323	competence protein CglA
ORF00180	SAG0164	282	competence protein CglB
ORF00181	SAG0165	151	conserved hypothetical protein
ORF00182	SAG0166	123	conserved domain protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00183	SAG0167	324	conserved hypothetical protein
ORF00184	SAG0168	397	acetate kinase
ORF00186	SAG0169	68	transcriptional regulator, Cro/CI family
ORF00187	SAG0170	45	hypothetical protein
ORF00188	SAG0171	151	hypothetical protein
ORF00189	SAG0172	221	protease, putative
ORF00190	SAG0173	256	pyrroline-5-carboxylate reductase
ORF00191	SAG0174	355	glutamyl-aminopeptidase
ORF00192	SAG0175	79	hypothetical protein
ORF00193	SAG0176	94	conserved hypothetical protein
ORF00194	SAG0177	107	thioredoxin family protein
ORF00195	SAG0178	208	tRNA binding domain protein
ORF00196	SAG0179	238	conserved hypothetical protein
ORF00198	SAG0180	131	single-strand binding protein
ORF00199	SAG0181	214	hydrolase, haloacid dehalogenase-like family
ORF00200	SAG0182	581	sensor histidine kinase, putative
ORF00201	SAG0183	246	response regulator
ORF00203	SAG0184	151	conserved hypothetical protein
ORF00204	SAG0185	242	membrane protein, putative
ORF00205	SAG0186	36	hypothetical protein
ORF00206	SAG0187	542	oligopeptide ABC transporter, oligopeptide-binding protein
ORF00207	SAG0188	325	oligopeptide ABC transporter, permease protein
ORF00208	SAG0189	273	oligopeptide ABC transporter, permease protein
ORF00209	SAG0190	267	peptide ABC transporter, ATP-binding protein
ORF00210	SAG0191	208	peptide ABC transporter, ATP-binding protein
ORF00211	SAG0192	676	PTS system, IIABC components
ORF00212	SAG0193	541	alpha amylase family protein
ORF00214	SAG0194	639	transcriptional antiterminator, BglG family
ORF00216	SAG0195	377	IS1548, transposase
ORF00217	SAG0196	66	conserved domain protein
ORF00218	SAG0197	94	PTS system, IIB component, putative
ORF00219	SAG0198	451	PTS system, IIC component, putative
ORF00220	SAG0199	285	transketolase, N-terminal subunit
ORF00221	SAG0200	309	transketolase, C-terminal subunit
ORF00223	SAG0201	419	oxidoreductase, putative
ORF00224	SAG0202	89	ribosomal protein S15
ORF00225	SAG0203	709	polyribonucleotide nucleotidyltransferase
ORF00226	SAG0204	250	conserved hypothetical protein
ORF00227	SAG0205	194	serine O-acetyltransferase
ORF00228	SAG0206	60	hypothetical protein
ORF00229	SAG0207	447	cysteinyl-tRNA synthetase
ORF00230	SAG0208	128	conserved hypothetical protein
ORF00231	SAG0209	251	RNA methyltransferase, TrmH family, group 3
ORF00232	SAG0210	172	conserved hypothetical protein
ORF00233	SAG0211	286	DegV family protein
ORF00234	SAG0212	32	hypothetical protein
ORF00235	SAG0213	39	hypothetical protein
ORF00236	SAG0214	148	ribosomal protein L13
ORF00237	SAG0215	130	ribosomal protein S9
ORF00238	SAG0216	33	hypothetical protein
ORF00239	SAG0217	384	site-specific recombinase, phage integrase family
ORF00240	SAG0218	158	transcriptional regulator, Cro/CI family
ORF00241	SAG0219	101	hypothetical protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00242	SAG0220	92	conserved hypothetical protein
ORF00243	SAG0221	76	hypothetical protein
ORF00244	SAG0222	108	conserved domain protein
ORF00245	SAG0223	209	conserved hypothetical protein, fusion
ORF00246	SAG0224	332	replication initiation protein, putative
ORF00247	SAG0225	144	hypothetical protein
ORF00248	SAG0226	418	recombination protein
ORF00249	SAG0227	156	hypothetical protein
ORF00250	SAG0228	111	conserved hypothetical protein
ORF00251	SAG0229	95	conserved hypothetical protein
ORF00252	SAG0230	96	conserved hypothetical protein
ORF00253	SAG0231	135	hypothetical protein
ORF00254	SAG0232	186	hypothetical protein
ORF00255	SAG0233	226	hypothetical protein
ORF00256	SAG0234	128	hypothetical protein
ORF00257	SAG0235	93	hypothetical protein
ORF00258	SAG0236	32	hypothetical protein
ORF00259	SAG0237	34	hypothetical protein
ORF00260	SAG0238	41	hypothetical protein
ORF00261	SAG0239	286	transcriptional regulator MutR family
ORF00262	SAG0240	393	transporter, putative
ORF00263	SAG0241	213	amino acid ABC transporter, permease protein
ORF00264	SAG0242	308	amino acid ABC transporter, amino acid-binding protein
ORF00265	SAG0243	211	amino acid ABC transporter, permease protein
ORF00266	SAG0244	381	amino acid ABC transporter, ATP-binding protein
ORF00272	SAG0245	152	hypothetical protein
ORF00273	SAG0246	268	hypothetical protein
ORF00274	SAG0247	116	hypothetical protein
ORF00275	SAG0248	90	hypothetical protein
ORF00276	SAG0249	116	hypothetical protein
ORF00278	SAG0250	193	hypothetical protein
ORF00279	SAG0251	72	transcriptional regulator, Cro/CI family
ORF00280	SAG0252	186	acetyltransferase, GNAT family
ORF00281	SAG0253	192	acetyltransferase, GNAT family
ORF00282	SAG0254	226	acetyltransferase, GNAT family
ORF00283	SAG0255	315	conserved hypothetical protein
ORF00284	SAG0256	163	RNA polymerase sigma factor, ECF subfamily
ORF00285	SAG0257	53	hypothetical protein
ORF00287	SAG0258	202	transcriptional regulator, TetR family
ORF00288	SAG0259	365	ABC transporter efflux protein, DrrB family, putative
ORF00289	SAG0260	238	ABC transporter, ATP-binding protein
ORF00290	SAG0261	129	IS1381, transposase OrfB
ORF00291	SAG0262	127	IS1381, transposase OrfA
ORF00292	SAG0263	171	hypothetical protein
ORF00293	SAG0264	103	conserved hypothetical protein
ORF00294	SAG0265	235	conserved hypothetical protein
ORF00295	SAG0266	382	N-acetylglucosamine-6-phosphate deacetylase
ORF00296	SAG0267	180	conserved hypothetical protein
ORF00297	SAG0268	304	glycyl-tRNA synthetase, alpha subunit
ORF00298	SAG0269	213	acyl carrier protein phosphodiesterase, putative
ORF00299	SAG0270	679	glycyl-tRNA synthetase, beta subunit
ORF00300	SAG0271	85	conserved hypothetical protein
ORF00301	SAG0272	87	membrane protein, putative

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00302	SAG0273	502	glycerol kinase
ORF00303	SAG0274	609	alpha-glycerophosphate oxidase
ORF00304	SAG0275	232	glycerol uptake facilitator protein
ORF00305	SAG0276	445	NADH oxidase, putative
ORF00306	SAG0277	476	conserved hypothetical protein
ORF00307	SAG0278	661	transketolase
ORF00308	SAG0279	101	conserved hypothetical protein
ORF00309	SAG0280	244	ABC transporter, ATP-binding protein
ORF00310	SAG0281	534	membrane protein, putative
ORF00313	SAG0282	461	PTS system, IIBC components
ORF00314	SAG0283	267	glutamate 5-kinase
ORF00315	SAG0284	417	gamma-glutamyl phosphate reductase
ORF00316	SAG0285	298	conserved hypothetical protein TIGR00006
ORF00317	SAG0286	108	cell division protein FtsL, putative
ORF00318	SAG0287	752	penicillin-binding protein 2X
ORF00319	SAG0288	336	phospho-N-acetylmuramoyl-pentapeptide-transferase
ORF00320	SAG0289	447	ATP-dependent RNA helicase, DEAD/DEAH box family
ORF00321	SAG0290	270	ABC transporter, substrate-binding protein
ORF00322	SAG0291	267	amino acid ABC transporter, permease protein
ORF00323	SAG0292	247	amino acid ABC transporter, ATP-binding protein
ORF00324	SAG0293	74	conserved hypothetical protein
ORF00325	SAG0294	304	thioredoxin reductase
ORF00326	SAG0295	486	conserved hypothetical protein
ORF00327	SAG0296	273	NAD synthetase
ORF00328	SAG0297	444	aminopeptidase C
ORF00329	SAG0298	750	penicillin-binding protein 1A
ORF00330	SAG0299	199	recombination protein U
ORF00331	SAG0300	172	conserved hypothetical protein
ORF00332	SAG0301	40	hypothetical protein
ORF00333	SAG0302	110	conserved hypothetical protein
ORF00335	SAG0303	384	conserved hypothetical protein
ORF00336	SAG0304	487	conserved hypothetical protein
ORF00337	SAG0305	160	autoinducer-2 production protein LuxS
ORF00338	SAG0306	535	KH domain protein
ORF00340	SAG0307	33	hypothetical protein
ORF00341	SAG0308		ABC transporter, ATP-binding protein, FRAMESHIFT
ORF00343	SAG0309	246	ABC transporter, permease protein, putative
ORF00344	SAG0310	361	conserved hypothetical protein
ORF00345	SAG0311		DNA-binding response regulator POINT MUTATION
ORF00347	SAG0312	234	conserved hypothetical protein
ORF00348	SAG0313	209	guanylate kinase
ORF00349	SAG0314	104	DNA-directed RNA polymerase, omega subunit, putative
ORF00350	SAG0315	796	primosomal protein N'
ORF00351	SAG0316	311	methionyl-tRNA formyltransferase
ORF00352	SAG0317	440	Sun protein
ORF00353	SAG0318	245	serine/threonine phosphatase, putative
ORF00354	SAG0319	651	serine/threonine protein kinase
ORF00355	SAG0320	231	conserved hypothetical protein
ORF00356	SAG0321	339	sensor histidine kinase, putative
ORF00358	SAG0322	213	DNA-binding response regulator

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00359	SAG0323	466	hydrolase, haloacid dehalogenase family/peptidyl-prolyl cis-trans isomerase, cyclophilin type
ORF00360	SAG0324	124	general stress protein, putative
ORF00361	SAG0325	258	pyruvate formate-lyase-activating enzyme
ORF00362	SAG0326	251	transcriptional regulator, DeoR family
ORF00363	SAG0327	327	transcriptional regulator, putative
ORF00364	SAG0328	107	PTS system, cellobiose-specific IIA component
ORF00366	SAG0329	106	PTS system, cellobiose-specific IIB component
ORF00367	SAG0330	433	PTS system, cellobiose-specific IIC component
ORF00368	SAG0331	818	formate acetyltransferase
ORF00369	SAG0332	222	transaldolase family protein
ORF00371	SAG0333	362	glycerol dehydrogenase
ORF00372	SAG0334	308	cysteine synthase A
ORF00373	SAG0335	214	conserved hypothetical protein TIGR00257
ORF00374	SAG0336	429	helicase, putative
ORF00375	SAG0337	221	competence protein F, putative
ORF00376	SAG0338	184	ribosomal subunit interface protein
ORF00382	SAG0339	450	aspartate kinase family protein
ORF00383	SAG0340	216	hydrolase, haloacid dehalogenase-like family
ORF00384	SAG0341	49	hypothetical protein
ORF00385	SAG0342	263	enoyl-CoA hydratase/isomerase family protein
ORF00386	SAG0343	144	transcriptional regulator, MarR family
ORF00387	SAG0344	323	3-oxoacyl-(acyl-carrier-protein) synthase III
ORF00388	SAG0345	74	acyl carrier protein
ORF00390	SAG0346	319	enoyl-(acyl-carrier-protein) reductase II
ORF00391	SAG0347	308	malonyl CoA-acyl carrier protein transacylase
ORF00392	SAG0348	244	3-oxoacyl-[acyl-carrier protein] reductase
ORF00393	SAG0349	410	3-oxoacyl-(acyl-carrier-protein) synthase II
ORF00394	SAG0350	166	acetyl-CoA carboxylase, biotin carboxyl carrier protein
ORF00395	SAG0351	140	(3R)-hydroxymyristoyl-(acyl-carrier-protein) dehydratase
ORF00396	SAG0352	456	acetyl-CoA carboxylase, biotin carboxylase
ORF00397	SAG0353	291	acetyl-CoA carboxylase, carboxyl transferase, beta subunit
ORF00398	SAG0354	257	acetyl-CoA carboxylase, carboxyl transferase, alpha subunit
ORF00399	SAG0355	210	conserved hypothetical protein
ORF00400	SAG0356	425	seryl-tRNA synthetase
ORF00402	SAG0357	330	hypothetical protein
ORF00403	SAG0358	120	conserved hypothetical protein
ORF00404	SAG0359	303	PTS system, mannose-specific IID component
ORF00405	SAG0360	270	PTS system, mannose-specific IIC component
ORF00406	SAG0361	336	PTS system, mannose-specific IIAB components
ORF00407	SAG0362	270	hydrolase, haloacid dehalogenase-like family
ORF00408	SAG0363	194	hypothetical protein
ORF00409	SAG0364	203	membrane protein, putative
ORF00410	SAG0365	473	xanthine/uracil permease family protein
ORF00411	SAG0366	169	conserved hypothetical protein TIGR00150, putative
ORF00412	SAG0367	186	acetyltransferase, GNAT family
ORF00413	SAG0368	435	transcriptional regulator, putative
ORF00414	SAG0369	98	conserved hypothetical protein
ORF00415	SAG0370	139	HIT family protein
ORF00416	SAG0371	167	hypothetical protein
ORF00417	SAG0372	85	hypothetical protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00419	SAG0373	241	ABC transporter, ATP-binding protein
ORF00421	SAG0374	344	ABC transporter, permease protein
ORF00422	SAG0375	266	conserved hypothetical protein
ORF00423	SAG0376	211	conserved hypothetical protein TIGR00091
ORF00424	SAG0377	127	conserved hypothetical protein, POINT MUTATION
ORF00425	SAG0378	379	N utilization substance protein A
ORF00426	SAG0379	98	conserved hypothetical protein
ORF00427	SAG0380	100	ribosomal protein L7A family
ORF00428	SAG0381	927	translation initiation factor IF-2
ORF00429	SAG0382	122	ribosome-binding factor A
ORF00430	SAG0383	334	conserved hypothetical protein
ORF00431	SAG0384	138	transcriptional repressor CopY
ORF00432	SAG0385	744	copper-transporter ATPase CopA
ORF00433	SAG0386	68	copper-transporter protein CopZ
ORF00434	SAG0387	204	conserved hypothetical protein
ORF00435	SAG0388	270	hydrolase, haloacid dehalogenase-like family
ORF00436	SAG0389	880	DNA polymerase I
ORF00437	SAG0390	146	CoA binding domain protein
ORF00438	SAG0391	159	transcriptional regulator, Fur family
ORF00439	SAG0392	521	cell wall surface anchor family protein
ORF00440	SAG0393	228	DNA-binding response regulator
ORF00441	SAG0394	345	sensor histidine kinase
ORF00442	SAG0395	246	conserved hypothetical protein
ORF00443	SAG0396	380	queuine tRNA-ribosyltransferase
ORF00444	SAG0397	102	conserved hypothetical protein
ORF00445	SAG0398	179	bioY family protein
ORF00446	SAG0399	258	AtsA/ElaC family protein
ORF00447	SAG0400	168	cytidine/deoxycytidylate deaminase family protein
ORF00448	SAG0401	44	hypothetical protein
ORF00449	SAG0402	449	glucose-6-phosphate isomerase
ORF00450	SAG0403	175	5-formyltetrahydrofolate cyclo-ligase family protein
ORF00451	SAG0404	225	rhomboid family protein
ORF00452	SAG0405	347	lipoprotein
ORF00453	SAG0406	299	UTP-glucose-1-phosphate uridylyltransferase
ORF00454	SAG0407	338	glycerol-3-phosphate dehydrogenase (NAD(P)+)
ORF00455	SAG0408	109	ribonuclease P protein component
ORF00456	SAG0409	271	SpoIIIJ family protein
ORF00458	SAG0410	273	R3H domain protein
ORF00463	SAG0411	177	conserved hypothetical protein
ORF00464	SAG0412	258	RecX protein
ORF00465	SAG0413	451	RNA methyltransferase, TrmA family
ORF00466	SAG0414	153	conserved hypothetical protein
ORF00467	SAG0415	142	acetyltransferase, GNAT family
ORF00468	SAG0416	1233	protease, putative
ORF00469	SAG0417	302	glycosyl transferase, group 2 family protein
ORF00470	SAG0418	336	ribonucleoside-diphosphate reductase 2, beta subunit
ORF00471	SAG0419	137	nrdI protein
ORF00472	SAG0420	721	ribonucleoside-diphosphate reductase 2, alpha subunit
ORF00473	SAG0421	1055	conserved hypothetical protein
ORF00474	SAG0422	129	conserved hypothetical protein
ORF00475	SAG0423	132	conserved domain protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00476	SAG0424	94	hypothetical protein
ORF00478	SAG0425	105	carboxymuconolactone decarboxylase family protein
ORF00479	SAG0426	131	conserved hypothetical protein
ORF00480	SAG0427	129	transcriptional regulator, MerR family
ORF00482	SAG0428	345	alcohol dehydrogenase, zinc-containing
ORF00483	SAG0429	284	oxidoreductase, aldo/keto reductase family
ORF00484	SAG0430	287	cation efflux system protein
ORF00485	SAG0431	174	transcriptional regulator, TetR family
ORF00486	SAG0432	397	transcriptional regulator, AraC family
ORF00487	SAG0433	1389	surface protein Rib
ORF00488	SAG0434	61	transposase, IS256 family, truncation
ORF00489	SAG0435	97	DNA-damage-inducible protein J, putative
ORF00490	SAG0436	62	hypothetical protein
ORF00491	SAG0437	123	hypothetical protein
ORF00493	SAG0438	145	bacteriophage L54a, integrase, truncation
ORF00495	SAG0439		conserved hypothetical protein, FRAMESHIFT
ORF00496	SAG0440	84	conserved hypothetical protein
ORF00497	SAG0441	103	conserved domain protein
ORF00499	SAG0442	189	acetyltransferase, GNAT family
ORF00500	SAG0443	194	acetyltransferase, GNAT family
ORF00501	SAG0444	188	conserved hypothetical protein
ORF00502	SAG0445	883	valyl-tRNA synthetase
ORF00503	SAG0446	319	oxidoreductase, Gfo/Idh/MocA family
ORF00504	SAG0447	287	magnesium transporter, CorA family
ORF00506	SAG0448	391	transposase, IS256 family
ORF00507	SAG0449	354	conserved hypothetical protein
ORF00508	SAG0450	330	aspartate--ammonia ligase
ORF00510	SAG0451	149	bacteriocin transport accessory protein,putative
ORF00511	SAG0452	179	type II DNA modification methyltransferase, putative
ORF00512	SAG0453	96	hypothetical protein
ORF00513	SAG0454	161	phosphopantetheine adenylyltransferase
ORF00515	SAG0455	357	conserved hypothetical protein
ORF00518	SAG0456		conserved hypothetical protein, degenerate
ORF00519	SAG0457	192	conserved hypothetical protein
ORF00520	SAG0458	368	conserved hypothetical protein TIGR00048
ORF00521	SAG0459	171	VanZF domain protein
ORF00522	SAG0460	581	ABC transporter, ATP-binding/permease protein
ORF00523	SAG0461	579	ABC transporter, ATP-binding/permease protein
ORF00524	SAG0462	188	anthranilate synthase component II
ORF00525	SAG0463	179	bioY family protein
ORF00526	SAG0464	330	biotin synthetase
ORF00527	SAG0465	164	hypothetical protein
ORF00528	SAG0466	371	thiolase
ORF00531	SAG0467	409	AMP-binding enzyme domain protein
ORF00532	SAG0468	210	endonuclease III
ORF00533	SAG0469	131	type IV prepilin peptidase-related protein
ORF00534	SAG0470	69	conserved hypothetical protein
ORF00535	SAG0471	322	glucokinase
ORF00536	SAG0472	126	rhodanese domain protein
ORF00537	SAG0473	613	elongation factor Tu family protein
ORF00538	SAG0474	81	conserved hypothetical protein
ORF00540	SAG0475	451	UDP-N-acetylmuramoylalanine--D-glutamate ligase

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00541	SAG0476	358	UDP-N-acetylglucosamine--N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase
ORF00542	SAG0477	378	cell division protein DivIB, putative
ORF00544	SAG0478	429	cell division protein FtsA
ORF00545	SAG0479	426	cell division protein FtsZ
ORF00546	SAG0480	224	ylmE protein, putative
ORF00547	SAG0481	201	ylmF protein
ORF00548	SAG0482	84	YGGT family protein
ORF00549	SAG0483	262	ylmH protein
ORF00550	SAG0484	256	cell division protein DivIVA, putative
ORF00552	SAG0485	930	isoleucyl-tRNA synthetase
ORF00553	SAG0486	100	conserved hypothetical protein
ORF00554	SAG0487	151	MutT/nudix family protein
ORF00555	SAG0488	753	ATP-dependent Clp protease, ATP-binding subunit
ORF00556	SAG0489	34	hypothetical protein
ORF00557	SAG0490	76	conserved hypothetical protein
ORF00558	SAG0491	230	amino acid ABC transporter, permease protein
ORF00559	SAG0492	244	amino acid ABC transporter, ATP-binding protein
ORF00560	SAG0493	564	phosphoglucomutase/phosphomannomutase family protein
ORF00562	SAG0494	284	methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase
ORF00563	SAG0495	278	conserved hypothetical protein
ORF00564	SAG0496	446	exodeoxyribonuclease VII, large subunit
ORF00565	SAG0497	71	exodeoxyribonuclease VII, small subunit
ORF00566	SAG0498	290	geranyltranstransferase, putative
ORF00567	SAG0499	275	hemolysin A
ORF00568	SAG0500	157	arginine repressor ArgR, putative
ORF00570	SAG0501	552	DNA repair protein RecN
ORF00571	SAG0502	278	DegV family protein
ORF00572	SAG0503	279	Lipase/Acylhydrolase, putative
ORF00573	SAG0504	200	conserved hypothetical protein
ORF00574	SAG0505	91	DNA-binding protein HU
ORF00575	SAG0506	65	hypothetical protein
ORF00576	SAG0507	310	dihydroorotate dehydrogenase A
ORF00577	SAG0508	411	beta-lactam resistance factor
ORF00578	SAG0509	403	beta-lactam resistance factor
ORF00579	SAG0510	406	murM protein, putative
ORF00580	SAG0511	270	hydrolase, haloacid dehalogenase-like family
ORF00581	SAG0512	438	HD domain protein
ORF00582	SAG0513	128	conserved hypothetical protein
ORF00583	SAG0514	894	cation-transporting ATPase, E1-E2 family
ORF00584	SAG0515	286	conserved hypothetical protein
ORF00585	SAG0516	643	fructose-1,6-bisphosphatase, putative
ORF00586	SAG0517	374	iron-sulfur cluster-binding protein, putative
ORF00587	SAG0518		peptide chain release factor 2, FRAMESHIFT
ORF00588	SAG0519	230	cell division ABC transporter, ATP-binding protein FtsE
ORF00589	SAG0520	309	cell division ABC transporter, permease protein FtsX
ORF00590	SAG0521	236	carboxymethylenebutenolidase-related protein
ORF00591	SAG0522	232	metallo-beta-lactamase superfamily protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00592	SAG0523	254	oxidoreductase, short chain dehydrogenase/reductase family
ORF00593	SAG0524	835	DNA polymerase III, epsilon subunit/ATP-dependent helicase DinG
ORF00595	SAG0525	397	aspartate aminotransferase
ORF00596	SAG0526	448	asparaginyl-tRNA synthetase
ORF00597	SAG0527	185	conserved hypothetical protein
ORF00598	SAG0528	327	inosine-uridine preferring nucleoside hydrolase
ORF00599	SAG0529	38	hypothetical protein
ORF00600	SAG0530	137	OsmC/Ohr family protein
ORF00601	SAG0531	296	conserved hypothetical protein
ORF00602	SAG0532	324	conserved hypothetical protein
ORF00603	SAG0533	303	Uncharacterized BCR, COG1481
ORF00604	SAG0534	465	dipeptidase
ORF00605	SAG0535	506	zinc ABC transporter, zinc-binding adhesion liprotein
ORF00606	SAG0536	86	ribosomal protein L31
ORF00607	SAG0537	311	DHH family protein
ORF00608	SAG0538	340	adenosine deaminase, putative
ORF00609	SAG0539	147	flavodoxin
ORF00610	SAG0540	91	chorismate mutase, putative
ORF00611	SAG0541	398	voltage-gated chloride channel family protein
ORF00612	SAG0542	127	IS1381, transposase OrfA
ORF00613	SAG0543	129	IS1381, transposase OrfB
ORF00614	SAG0544	115	ribosomal protein L19
ORF00615	SAG0545	359	site-specific recombinase, phage integrase family
ORF00617	SAG0546	67	conserved domain protein
ORF00618	SAG0547	185	hypothetical protein
ORF00619	SAG0548	265	repressor protein, putative
ORF00620	SAG0549	47	hypothetical protein
ORF00621	SAG0550	74	conserved hypothetical protein
ORF00622	SAG0551	52	conserved hypothetical protein
ORF00623	SAG0552	62	hypothetical protein
ORF00624	SAG0553	268	hypothetical protein
ORF00626	SAG0554	63	transcriptional regulator, Cro/CI family
ORF00627	SAG0555	249	antirepressor, putative
ORF00628	SAG0556	47	hypothetical protein
ORF00630	SAG0557	76	hypothetical protein
ORF00632	SAG0558	74	hypothetical protein
ORF00633	SAG0559	286	conserved hypothetical protein
ORF00634	SAG0560	77	conserved hypothetical protein
ORF00635	SAG0561	46	hypothetical protein
ORF00636	SAG0562	84	hypothetical protein
ORF00637	SAG0563	53	hypothetical protein
ORF00638	SAG0564	160	conserved hypothetical protein
ORF00639	SAG0565	224	conserved domain protein
ORF00640	SAG0566	138	single-strand binding protein
ORF00641	SAG0567	439	reverse transcriptase/maturase family protein
ORF00642	SAG0568	67	conserved hypothetical protein
ORF00643	SAG0569	158	conserved hypothetical protein
ORF00644	SAG0570	115	hypothetical protein
ORF00645	SAG0571	43	hypothetical protein
ORF00646	SAG0572	138	conserved hypothetical protein
ORF00647	SAG0573	54	hypothetical protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00648	SAG0574	89	conserved hypothetical protein
ORF00649	SAG0575	110	hypothetical protein
ORF00650	SAG0576	43	hypothetical protein
ORF00652	SAG0577	177	conserved hypothetical protein
ORF00653	SAG0578	88	conserved hypothetical protein
ORF00654	SAG0581	118	conserved hypothetical protein
ORF00655	SAG0582	422	conserved hypothetical protein
ORF00656	SAG0583	406	conserved hypothetical protein
ORF00657	SAG0584	62	conserved hypothetical protein, truncation
ORF00658	SAG0585	471	conserved hypothetical protein
ORF00659	SAG0586	154	conserved hypothetical protein
ORF00660	SAG0587	300	structural protein, putative
ORF00661	SAG0588	71	conserved hypothetical protein
ORF00662	SAG0589	143	conserved hypothetical protein
ORF00663	SAG0590	112	conserved hypothetical protein
ORF00664	SAG0591	78	conserved hypothetical protein
ORF00665	SAG0592	111	conserved hypothetical protein
ORF00666	SAG0593	185	structural protein
ORF00667	SAG0594	81	conserved hypothetical protein
ORF00668	SAG0595	123	conserved hypothetical protein
ORF00669	SAG0596	670	PblA, internal deletion
ORF00670	SAG0597	506	minor structural protein, putative
ORF00671	SAG0598	1374	minor structural protein, putative
ORF00672	SAG0599	668	minor structural protein, putative
ORF00673	SAG0600	109	hypothetical protein
ORF00674	SAG0601	70	hypothetical protein
ORF00675	SAG0602	100	conserved hypothetical protein
ORF00676	SAG0603	111	holin, putative
ORF00677	SAG0604	239	lysin, putative
ORF00678	SAG0605	323	conserved hypothetical protein
ORF00679	SAG0606	66	conserved hypothetical protein
ORF00681	SAG0607	56	conserved hypothetical protein
ORF00682	SAG0608	59	hypothetical protein
ORF00683	SAG0609	193	site-specific recombinase, phage integrase family
ORF00685	SAG0610	134	conserved hypothetical protein
ORF00687	SAG0611		transposase, degenerate FRAMESHIFT
ORF00689	SAG0612	53	conserved hypothetical protein, FRAMESHIFT
ORF00690	SAG0613	425	transmembrane protein Vexp1
ORF00691	SAG0614	218	ABC transporter, ATP-binding protein Vexp2
ORF00692	SAG0615	458	transmembrane protein Vexp3
ORF00693	SAG0616	217	DNA-binding response regulator VncR
ORF00694	SAG0617	439	sensor histidine kinase VncS
ORF00695	SAG0618	195	transposase OrfB, IS3 family, truncation
ORF00697	SAG0619	66	conserved hypothetical protein
ORF00698	SAG0620	62	hypothetical protein
ORF00699	SAG0621	401	rod shape-determining protein RodA, putative□
ORF00700	SAG0622	186	hydrolase, haloacid dehalogenase-like family
ORF00701	SAG0623	650	DNA gyrase, B subunit
ORF00702	SAG0624	574	septation ring formation regulator EzrA, putative
ORF00703	SAG0625	213	phosphoserine phosphatase SerB
ORF00704	SAG0626	161	MutT/nudix family protein
ORF00705	SAG0627	151	conserved hypothetical protein
ORF00706	SAG0628	435	enolase
ORF00707	SAG0629	354	conserved domain protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00708	SAG0630	427	3-phosphoshikimate 1-carboxyvinyltransferase
ORF00709	SAG0631	170	shikimate kinase
ORF00710	SAG0632	457	psr protein
ORF00711	SAG0633	451	RNA methyltransferase, TrmA family
ORF00712	SAG0634	70	hypothetical protein
ORF00713	SAG0635	245	acid phosphatase precursor, class B
ORF00714	SAG0636	172	conserved hypothetical protein
ORF00717	SAG0637		transcriptional regulator, TetR family, putative, FRAMESHIFT
ORF00718	SAG0638	109	cell wall surface anchor family protein
ORF00720	SAG0639	273	transposase OrfB, IS3 family
ORF00721	SAG0640	91	transposase OrfA, IS3 family
ORF00722	SAG0641		Tn5252, Orf 10 protein, degenerate POINT MUTATION
ORF00723	SAG0642	59	hypothetical protein
ORF00725	SAG0643		chaperonin, 33 kDa DEGENERATE
ORF00726	SAG0644	402	transcriptional regulator, AraC family
ORF00727	SAG0645	554	cell wall surface anchor family protein, putative
ORF00728	SAG0646	307	cell wall surface anchor family protein
ORF00729	SAG0647	305	sortase family protein
ORF00731	SAG0648	260	sortase family protein
ORF00732	SAG0649	890	cell wall surface anchor family protein,putative
ORF00734	SAG0650	189	sortase family protein, FRAMESHIFT
ORF00735	SAG0651	201	hypothetical protein
ORF00737	SAG0653	76	conserved hypothetical protein, DEGENERATE
ORF00738	SAG0654	34	hypothetical protein
ORF00740	SAG0656	36	hypothetical protein
ORF00741	SAG0657	89	hypothetical protein
ORF00742	SAG0658	383	lipoprotein, putative
ORF00743	SAG0659	330	ABC transporter, ATP-binding protein
ORF00744	SAG0660	272	membrane protein
ORF00745	SAG0661	261	conserved hypothetical protein
ORF00747	SAG0663	282	cylD protein
ORF00748	SAG0664	240	cylG protein
ORF00749	SAG0665	101	acyl carrier protein AcpC
ORF00750	SAG0666	158	cylZ protein FRAMESHIFT
ORF00751	SAG0667	309	cylA protein
ORF00752	SAG0668	292	cylB protein
ORF00753	SAG0669	667	cylE protein
ORF00754	SAG0670	317	cylF protein
ORF00755	SAG0671	731	cylI protein
ORF00756	SAG0672	403	cylJ protein
ORF00757	SAG0673	191	cylK protein
ORF00758	SAG0674	113	hypothetical protein
ORF00759	SAG0675	171	surface protein antigen-related protein
ORF00760	SAG0676	885	serine protease, subtilase family, putative
ORF00761	SAG0677	1062	hypothetical protein
ORF00762	SAG0678		endopeptidase O DEGENERATE
ORF00766	SAG0679	286	hydrolase, alpha/beta fold family, putative
ORF00767	SAG0680	339	hypothetical protein
ORF00768	SAG0681	353	conserved domain protein
ORF00769	SAG0682	409	permease, putative
ORF00770	SAG0683		transmembrane protein Vexp3, putative FRAMESHIFT
ORF00774	SAG0684	223	ABC transporter, ATP-binding protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00775	SAG0685	472	conserved hypothetical protein
ORF00776	SAG0686	261	DNA-entry nuclease, putative
ORF00777	SAG0687	212	DedA family protein, putative
ORF00778	SAG0688	218	ABC transporter, ATP-binding protein
ORF00779	SAG0689	257	membrane protein, putative
ORF00780	SAG0690	272	conserved hypothetical protein
ORF00781	SAG0691	294	transcriptional regulator, LysR family
ORF00783	SAG0692	193	regulatory protein, putative
ORF00785	SAG0693	377	IS1548, transposase
ORF00786	SAG0694	173	regulatory protein, putative, truncation
ORF00787	SAG0695	330	D-lactate dehydrogenase
ORF00788	SAG0696	516	sodium:galactoside symporter family protein, putative
ORF00789	SAG0697	341	2-keto-3-deoxygluconate kinase
ORF00790	SAG0698	599	beta-glucuronidase
ORF00791	SAG0699	223	transcriptional regulator, GntR family
ORF00792	SAG0700	205	2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase
ORF00793	SAG0701	466	Glucuronate isomerase
ORF00794	SAG0702	348	mannonate dehydratase
ORF00795	SAG0703	279	D-mannonate oxidoreductase
ORF00796	SAG0704	270	hydrolase, haloacid dehalogenase-like family
ORF00797	SAG0705	596	glycosyl hydrolase, family 3
ORF00798	SAG0706	361	proline dipeptidase
ORF00799	SAG0707	334	transcriptional regulator, RegM family
ORF00800	SAG0708	488	alpha amylase family protein
ORF00801	SAG0709	332	glycosyl transferase, group 1 family protein
ORF00802	SAG0710	444	glycosyl transferase, group 1 family protein
ORF00803	SAG0711	647	threonyl-tRNA synthetase
ORF00804	SAG0712	234	DNA-binding response regulator
ORF00805	SAG0713	339	conserved hypothetical protein
ORF00806	SAG0714	188	conserved hypothetical protein
ORF00807	SAG0715	216	amino acid ABC transporter, permease protein
ORF00808	SAG0716	231	amino acid ABC transporter, permease protein
ORF00809	SAG0717	266	amino acid ABC transporter, amino acid-binding protein
ORF00810	SAG0718	251	amino acid ABC transporter, ATP-binding protein
ORF00811	SAG0719	236	DNA-binding response regulator
ORF00812	SAG0720	449	sensory box histidine kinase
ORF00813	SAG0721	269	metallo-beta-lactamase family protein
ORF00814	SAG0722	122	conserved hypothetical protein
ORF00815	SAG0723	236	ribonuclease III
ORF00816	SAG0724	1179	SMC family protein
ORF00817	SAG0725	265	hydrolase, haloacid dehalogenase-like family
ORF00818	SAG0726	274	hydrolase, haloacid dehalogenase-like family
ORF00819	SAG0727	536	signal recognition particle-docking protein FtsY
ORF00820	SAG0728	270	ABC transporter, substrate-binding protein
ORF00821	SAG0729	300	ABC transporter, permease protein, putative
ORF00822	SAG0730	42	ABC transporter, ATP-binding protein
ORF00823	SAG0731	347	bacterial luciferase family protein
ORF00824	SAG0732	720	transcriptional accessory protein Tex, putative
ORF00825	SAG0733	142	conserved hypothetical protein
ORF00826	SAG0734	87	phage shock protein C, putative
ORF00827	**SAG0735**	**44**	**hypothetical protein**
ORF00828	SAG0736	311	HPr(Ser) kinase/phosphatase

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00830	SAG0737	257	prolipoprotein diacylglyceryl transferase
ORF00832	SAG0738	132	conserved hypothetical protein
ORF00833	SAG0739	143	conserved hypothetical protein
ORF00834	SAG0740	91	conserved hypothetical protein
ORF00835	SAG0741	303	peptidase, U32 family, putative
ORF00836	SAG0742	428	peptidase, U32 family
ORF00837	SAG0743	70	conserved hypothetical protein
ORF00838	SAG0744	265	membrane protein, putative
ORF00839	SAG0745	446	Mn2+/Fe2+ transporter, NRAMP family
ORF00840	SAG0746	369	riboflavin biosynthesis protein RibD
ORF00841	SAG0747	208	riboflavin synthase, alpha subunit
ORF00842	SAG0748	397	riboflavin biosynthesis protein RibA
ORF00843	SAG0749	156	riboflavin synthase, beta subunit
ORF00844	SAG0750	496	lysyl-tRNA synthetase
ORF00845	SAG0751	300	hydrolase, haloacid dehalogenase-like family
ORF00846	SAG0752	213	phosphoglycerate mutase family protein
ORF00847	SAG0753	157	ebsC family protein, putative
ORF00848	SAG0754	205	conserved domain protein
ORF00850	SAG0755	282	peptidase, U32 family
ORF00852	SAG0756	174	conserved hypothetical protein
ORF00853	SAG0757	129	lipoprotein, putative
ORF00855	SAG0758	599	oligoendopeptidase F, putative
ORF00856	SAG0759	931	phosphoenolpyruvate carboxylase
ORF00857	SAG0760	377	IS1548, transposase
ORF00859	SAG0761	422	cell division protein, FtsW/RodA/SpoVE family
ORF00861	SAG0762	398	translation elongation factor Tu
ORF00863	SAG0763	252	triosephosphate isomerase
ORF00865	SAG0764	230	phosphoglycerate mutase
ORF00866	SAG0765	681	penicillin-binding protein 2b
ORF00867	SAG0766	198	recombination protein RecR
ORF00868	SAG0767	348	D-alanine--D-alanine ligase
ORF00869	SAG0768	455	UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate--D-alanyl-D-alanyl ligase
ORF00870	SAG0769	406	oxalate:formate antiporter
ORF00871	SAG0770	228	conserved hypothetical protein
ORF00872	SAG0771	512	cell wall surface anchor family protein
ORF00873	SAG0772	514	peptide chain release factor 3
ORF00874	SAG0773	126	conserved hypothetical protein
ORF00876	SAG0774	244	ABC transporter, ATP-binding protein
ORF00878	SAG0775	220	ABC transporter, permease protein
ORF00879	SAG0776	276	lipoprotein,putative
ORF00880	SAG0777	528	ATP-dependent RNA helicase, DEAD/DEAH box family
ORF00882	SAG0778	88	conserved hypothetical protein
ORF00883	SAG0779	254	conserved hypothetical protein
ORF00884	SAG0780	246	acyltransferase family protein
ORF00885	SAG0781	217	competence protein CelA
ORF00887	SAG0782	745	DNA internalization-related competence protein ComEC/Rec2
ORF00888	SAG0783	269	hydrolase, haloacid dehalogenase-like family
ORF00889	SAG0784	314	sugar-binding transcriptional regulator, LacI family
ORF00890	SAG0785	330	conserved hypothetical protein
ORF00891	SAG0786	242	conserved domain protein
ORF00892	SAG0787	345	DNA polymerase III, delta subunit, putative□

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00893	SAG0788	202	superoxide dismutase, Fe-Mn
ORF00894	SAG0789	283	transcriptional antiterminator LicT
ORF00895	SAG0790	622	PTS system, beta-glucosides-specific IIABC components
ORF00896	SAG0791	475	6-phospho-beta-glucosidase
ORF00898	SAG0792	364	conserved hypothetical protein
ORF00899	SAG0793	380	conserved hypothetical protein TIGR00045
ORF00900	SAG0794	418	permease, GntP family
ORF00902	SAG0795	354	conserved hypothetical protein
ORF00903	SAG0796	147	transcriptional regulator, MarR family
ORF00904	SAG0797	342	S-adenosylmethionine:tRNA ribosyltransferase-isomerase
ORF00905	SAG0798	226	membrane protein, putative
ORF00906	SAG0799	233	glucosamine-6-phosphate isomerase
ORF00907	SAG0800	318	Glutathione S-transferases domain protein
ORF00908	SAG0801	239	ribosomal small subunit pseudouridine synthase
ORF00909	SAG0802	38	hypothetical protein
ORF00910	SAG0803	383	major facilitator family protein
ORF00911	SAG0804	315	competence protein CoiA
ORF00912	SAG0805	601	oligoendopeptidase B
ORF00913	SAG0806	208	hydrolase, haloacid dehalogenase-like family
ORF00914	SAG0807	235	O-methyltransferase family protein
ORF00916	SAG0808	309	protease maturation protein, putative
ORF00918	SAG0809	161	conserved hypothetical protein
ORF00919	SAG0810	872	alanyl-tRNA synthetase
ORF00921	SAG0811	238	membrane protein, putative
ORF00922	SAG0812	272	glycosyl transferase, family 8
ORF00923	SAG0813	81	hypothetical protein
ORF00924	SAG0814	95	conserved domain protein
ORF00925	SAG0815	71	transcriptional regulator, Cro/CI family
ORF00926	SAG0816	253	conserved hypothetical protein
ORF00927	SAG0817	187	conserved hypothetical protein
ORF00928	SAG0818	319	ribonucleoside-diphosphate reductase 2, beta subunit
ORF00929	SAG0819	719	ribonucleoside-diphosphate reductase 2, alpha subunit
ORF00930	SAG0820	74	ribonucleoside-diphosphate reductase 2, NrdH-redoxin
ORF00931	SAG0821	87	phosphocarrier protein HPr
ORF00932	SAG0822	577	phosphoenolpyruvate-protein phosphotransferase
ORF00933	SAG0823	475	glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent
ORF00934	SAG0824	417	polysaccharide deacetylase family protein
ORF00935	SAG0825	360	ATP-dependent RNA helicase, DEAD/DEAH box family
ORF00936	SAG0826	209	uridine kinase
ORF00937	SAG0827	165	conserved hypothetical protein
ORF00938	SAG0828	554	DNA polymerase III, gamma and tau subunits
ORF00939	SAG0829	64	conserved hypothetical protein
ORF00940	SAG0830	311	biotin--acetyl-CoA-carboxylase ligase
ORF00941	SAG0831	398	S-adenosylmethionine synthetase
ORF00942	SAG0832	753	hypothetical protein
ORF00943	SAG0833	181	hypothetical protein
ORF00944	SAG0834	42	hypothetical protein
ORF00945	SAG0835	188	conserved hypothetical protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00946	SAG0836	184	conserved hypothetical protein
ORF00948	SAG0837	428	ABC transporter, ATP-binding protein
ORF00950	SAG0838	233	hypothetical protein
ORF00951	SAG0839	226	transcriptional regulator, TenA family
ORF00952	SAG0840	265	phosphomethylpyrimidine kinase
ORF00953	SAG0841	256	hydroxyethylthiazole kinase
ORF00954	SAG0842	223	thiamine-phosphate pyrophosphorylase
ORF00955	SAG0843	419	UDP-N-acetylglucosamine 1-carboxyvinyltransferase
ORF00956	SAG0844	184	acetyltransferase, GNAT family
ORF00957	SAG0845	427	CBS domain protein
ORF00958	SAG0846	286	methionine aminopeptidase, type I
ORF00959	SAG0847	306	ribonuclease BN, putative
ORF00961	SAG0848	151	GtrA family protein
ORF00962	SAG0849	169	conserved hypothetical protein
ORF00963	SAG0850	652	DNA ligase, NAD-dependent
ORF00964	SAG0851	339	BmrU protein, putative
ORF00966	SAG0852	766	pullulanase, putative
ORF00967	SAG0853	622	1,4-alpha-glucan branching enzyme
ORF00968	SAG0854	379	glucose-1-phosphate adenylyltransferase
ORF00969	SAG0855		glycogen biosynthesis protein GlgD FRAMESHIFT
ORF00971	SAG0856	476	glycogen synthase
ORF00972	SAG0857	66	ATP synthase F0, C subunit
ORF00973	SAG0858	238	ATP synthase F0, A subunit
ORF00974	SAG0859	165	ATP synthase F0, B subunit
ORF00975	SAG0860	178	ATP synthase F1, delta subunit
ORF00976	SAG0861	501	ATP synthase F1, alpha subunit
ORF00977	SAG0862	293	ATP synthase F1, gamma subunit
ORF00978	SAG0863	468	ATP synthase F1, beta subunit
ORF00979	SAG0864	137	ATP synthase F1, epsilon subunit
ORF00980	SAG0865	76	conserved hypothetical protein
ORF00981	SAG0866	423	UDP-N-acetylglucosamine 1-carboxyvinyltransferase
ORF00982	SAG0867	63	conserved hypothetical protein
ORF00983	SAG0868	285	DNA-entry nuclease
ORF00984	SAG0869	346	phenylalanyl-tRNA synthetase, alpha subunit
ORF00985	SAG0870	173	acetyltransferase, GNAT family
ORF00986	SAG0871	801	phenylalanyl-tRNA synthetase, beta subunit
ORF00987	SAG0872	300	conserved hypothetical protein
ORF00988	SAG0873	1077	exonuclease RexB
ORF00989	SAG0874	1207	exonuclease RexA
ORF00990	SAG0875	305	magnesium transporter, CorA family, putative
ORF00991	SAG0876	458	tRNA modification GTPase TrmE
ORF00992	SAG0877	636	ABC transporter, ATP-binding protein
ORF00993	SAG0878	322	acetoin dehydrogenase, thymine PPi dependent, E1 component, alpha subunit
ORF00994	SAG0879	332	acetoin dehydrogenase, thymine PPi dependent, E1 component, beta subunit
ORF00995	SAG0880	462	acetoin dehydrogenase, thymine PPi dependent, E2 component, dihydrolipoamide acetyltransferase
ORF00996	SAG0881	585	acetoin dehydrogenase, thymine PPi dependent, E3 component, dihydrolipoamide dehydrogenase
ORF00997	SAG0882	329	lipoate-protein ligase A
ORF00998	SAG0883	261	cobyric acid synthase, putative

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF00999	SAG0884	447	mur ligase family protein
ORF01000	SAG0885	283	conserved hypothetical protein TIGR00159
ORF01001	SAG0886	319	Gram-positive signal peptide, YSIRK family domain protein
ORF01002	SAG0887	450	phosphoglucomutase/phosphomannomutase family protein
ORF01003	SAG0888	123	conserved hypothetical protein
ORF01004	SAG0889	126	conserved hypothetical protein
ORF01005	SAG0890	376	oxygen-independent coproporphyrinogen III oxidase, putative
ORF01006	SAG0891	245	conserved hypothetical protein
ORF01007	SAG0892	256	hydrolase, haloacid dehalogenase-like family
ORF01008	SAG0893	218	conserved hypothetical protein
ORF01009	SAG0894	1370	conserved hypothetical protein
ORF01010	SAG0895	289	lipoyl-binding domain protein
ORF01011	SAG0896	108	oxidoreductase, putative
ORF01012	SAG0897	221	conserved hypothetical protein
ORF01013	SAG0898	83	hypothetical protein
ORF01014	SAG0899	57	hypothetical protein
ORF01015	SAG0900	56	hypothetical protein
ORF01016	SAG0901	127	hypothetical protein
ORF01018	SAG0902	45	hypothetical protein
ORF01019	SAG0903	44	hypothetical protein
ORF01021	SAG0904	56	hypothetical protein
ORF01022	SAG0905	138	nucleoside diphosphate kinase
ORF01023	SAG0906	610	GTP-binding protein LepA
ORF01024	SAG0907	877	streptococcal histidine triad family protein
ORF01025	SAG0908	203	HD domain protein
ORF01026	SAG0909	154	acetyltransferase, GNAT family
ORF01027	SAG0910	144	PilB-related protein
ORF01030	SAG0911	930	cation-transporting ATPase, E1-E2 family
ORF01031	SAG0912	367	nucleoside diphosphate kinase domain protein
ORF01032	SAG0913	212	chloramphenicol acetyltransferase
ORF01033	SAG0914	203	conserved hypothetical protein
ORF01034	SAG0915	405	Tn916, transposase
ORF01035	SAG0916	67	Tn916, excisionase
ORF01037	SAG0918	76	Tn916, hypothetical protein
ORF01038	SAG0919	157	Tn916, hypothetical protein
ORF01039	SAG0921	117	Tn916, transcriptional regulator, putative
ORF01040	SAG0923	639	Tn916, tetracycline resistance protein
ORF01041	SAG0925	310	Tn916, hypothetical protein
ORF01042	SAG0926	333	Tn916, NLP/P60 family protein
ORF01044	SAG0927	725	Tn916, hypothetical protein FRAMESHIFT
ORF01047	SAG0928		Tn916, hypothetical protein FRAMESHIFT
ORF01048	SAG0929	168	Tn916, hypothetical protein
ORF01049	SAG0930	165	Tn916, hypothetical protein
ORF01050	SAG0931	73	Tn916, hypothetical protein
ORF01051	SAG0932	401	Tn916, transcriptional regulator, putative
ORF01052	SAG0933	461	Tn916, FtsK/SpoIIIE family protein
ORF01053	SAG0934	128	Tn916, hypothetical protein
ORF01054	SAG0935	104	Tn916, hypothetical protein
ORF01056	SAG0937		ABC transporter, ATP-binding protein, FRAMESHIFT
ORF01057	SAG0938	122	transcriptional regulator, GntR family
ORF01058	SAG0939	1034	DNA polymerase III, alpha subunit

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01059	SAG0940	340	6-phosphofructokinase
ORF01060	SAG0941	500	pyruvate kinase
ORF01061	SAG0942	185	signal peptidase I, putative
ORF01062	SAG0943	47	hypothetical protein
ORF01063	SAG0944	604	glucosamine--fructose-6-phosphate aminotransferase (isomerizing)
ORF01064	SAG0945	377	IS1548, transposase
ORF01066	SAG0946	109	phnA protein
ORF01068	SAG0947	213	amino acid ABC transporter, permease protein
ORF01069	SAG0948	209	amino acid ABC transporter, ATP-binding protein
ORF01070	SAG0949	276	amino acid ABC transporter, amino acid-binding protein
ORF01072	SAG0950	82	ribosomal protein S20
ORF01073	SAG0951	306	pantothenate kinase
ORF01074	SAG0952	196	conserved hypothetical protein
ORF01075	SAG0953	129	cytidine deaminase
ORF01076	SAG0954	349	lipoprotein
ORF01077	SAG0955	511	sugar ABC transporter, ATP-binding protein
ORF01078	SAG0956	353	sugar ABC transporter, permease protein, putative
ORF01079	SAG0957	318	sugar ABC transporter, permease protein, putative
ORF01080	SAG0958	456	NADH oxidase
ORF01081	SAG0959	329	L-lactate dehydrogenase
ORF01082	SAG0960	819	DNA gyrase, A subunit
ORF01083	SAG0961	247	sortase SrtA
ORF01084	SAG0962	137	glyoxylase family protein
ORF01085	SAG0963	320	conserved hypothetical protein
ORF01086	SAG0964	375	Na+/H+ exchanger family protein
ORF01087	SAG0965	127	IS1381, transposase OrfA
ORF01088	SAG0966	129	IS1381, transposase OrfB
ORF01089	SAG0967	520	GMP synthase
ORF01090	SAG0968	232	transcriptional regulator, GntR family
ORF01091	SAG0969	444	gid protein
ORF01092	SAG0970	247	acetyltransferase, GNAT family
ORF01093	SAG0971	282	lipoprotein,putative
ORF01095	SAG0972		conserved hypothetical protein, FRAMESHIFT
ORF01096	SAG0973	320	nisin-resistance protein, putative
ORF01097	SAG0974	250	ABC transporter, ATP-binding protein
ORF01098	SAG0975	651	ABC transporter, permease protein, putative
ORF01099	SAG0976	222	DNA-binding response regulator
ORF01100	SAG0977	312	sensor histidine kinase
ORF01101	SAG0978	356	site-specific recombinase, phage integrase family
ORF01102	SAG0979	553	ABC transporter, substrate binding protein, putative
ORF01103	SAG0980	257	conserved hypothetical protein
ORF01104	SAG0981	228	SatD
ORF01106	SAG0982	521	signal recognition particle protein
ORF01108	SAG0983	110	conserved hypothetical protein
ORF01109	SAG0984	437	sensor histidine kinase CiaH
ORF01110	SAG0985	226	DNA-binding response regulator CiaR
ORF01111	SAG0986	849	aminopeptidase N
ORF01112	SAG0987	217	phosphate transport system regulatory protein PhoU

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01113	SAG0988	252	phosphate ABC transporter, ATP-binding protein PstB, putative
ORF01114	SAG0989	267	phosphate ABC transporter, ATP-binding protein PstB, putative
ORF01115	SAG0990	295	phosphate ABC transporter, permease protein PstA, putative
ORF01116	SAG0991	305	phosphate ABC transporter, permease protein
ORF01117	SAG0992	286	phosphate ABC transporter, phosphate-binding protein
ORF01118	SAG0993	436	NOL1/NOP2/sun family protein
ORF01119	SAG0994	254	inositol monophosphatase family protein
ORF01120	SAG0995	93	conserved hypothetical protein
ORF01121	SAG0996	137	conserved hypothetical protein
ORF01122	SAG0997	310	macrolide-efflux protein mreA/riboflavin biosynthesis protein RibF
ORF01123	SAG0998	294	tRNA pseudouridine synthase B
ORF01124	SAG0999	143	acetyltransferase, GNAT family
ORF01125	SAG1000	423	conserved hypothetical protein
ORF01126	SAG1001	196	conserved hypothetical protein
ORF01127	SAG1002	292	protease, putative
ORF01128	SAG1003	876	permease, putative
ORF01129	SAG1004	233	ABC transporter, ATP-binding protein
ORF01131	SAG1005	706	DNA topoisomerase I
ORF01132	SAG1006	280	DprA/SMF protein, putative DNA processing factor
ORF01133	SAG1007	342	iron-compound ABC transporter, iron-compound-binding protein
ORF01134	SAG1008	253	iron compound ABC transporter, ATP-binding protein
ORF01135	SAG1009	324	iron compound ABC transporter, permease protein
ORF01136	SAG1010	320	iron compound ABC transporter, permease protein
ORF01137	SAG1011	182	acetyltransferase, CysE/LacA/LpxA/NodL family
ORF01138	SAG1012	253	ribonuclease HII
ORF01139	SAG1013	283	GTP-binding protein
ORF01140	SAG1014	190	conserved hypothetical protein
ORF01142	SAG1015	494	carbon starvation protein CstA, putative
ORF01143	SAG1016	244	response regulator
ORF01144	SAG1017	579	sensor histidine kinase, putative
ORF01145	SAG1018	40	hypothetical protein
ORF01146	SAG1019	39	conserved hypothetical protein, FRAMESHIFT
ORF01148	SAG1020	227	hypothetical protein
ORF01149	SAG1021	107	hypothetical protein
ORF01150	SAG1022	177	hypothetical protein
ORF01151	SAG1023	48	hypothetical protein
ORF01152	SAG1024	183	hypothetical protein
ORF01153	SAG1025	149	hypothetical protein
ORF01156	SAG1026		immunogenic secreted protein, DEGENERATE
ORF01157	SAG1027	84	conserved hypothetical protein
ORF01158	SAG1028	196	hypothetical protein
ORF01159	SAG1029	101	hypothetical protein
ORF01160	SAG1030	304	conserved hypothetical protein
ORF01161	SAG1031	120	extracellular protein, putative POINT MUATION
ORF01162	SAG1032	85	conserved hypothetical protein
ORF01164	SAG1033	1309	FtsK/SpoIIIE family protein
ORF01166	SAG1034	55	hypothetical protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01167	SAG1035	424	conserved hypothetical protein
ORF01168	SAG1036	80	conserved hypothetical protein
ORF01169	SAG1037	157	hypothetical protein
ORF01172	SAG1038	1003	phage infection protein, putative
ORF01173	SAG1039	96	conserved hypothetical protein
ORF01174	SAG1040	260	conserved domain protein
ORF01175	SAG1041	107	hypothetical protein
ORF01176	SAG1042	1060	carbamoyl-phosphate synthase, large subunit
ORF01177	SAG1043	358	carbamoyl-phosphate synthase, small subunit
ORF01178	SAG1044	307	aspartate carbamoyltransferase
ORF01179	SAG1045	430	dihydroorotase, multifunctional complex type
ORF01180	SAG1046	209	orotate phosphoribosyltransferase
ORF01181	SAG1047	233	orotidine 5'-phosphate decarboxylase
ORF01182	SAG1048	410	membrane protein, putative
ORF01183	SAG1049	513	ABC transporter, ATP-binding protein
ORF01184	SAG1050	112	ribonucleotide reductase, truncation
ORF01185	SAG1051	358	aspartate-semialdehyde dehydrogenase
ORF01186	SAG1052	47	cell wall surface anchor family protein, putative
ORF01187	SAG1053	30	hypothetical protein
ORF01188	SAG1054	531	cardiolipin synthetase
ORF01189	SAG1055	556	formate--tetrahydrofolate ligase
ORF01190	SAG1056	339	lipoate-protein ligase A
ORF01191	SAG1057	292	conserved hypothetical protein
ORF01192	SAG1058	272	conserved hypothetical protein
ORF01193	SAG1059	110	glycine cleavage system H protein, putative
ORF01194	SAG1060	328	bacterial luciferase family protein
ORF01195	SAG1061	399	oxidoreductase, FMN-binding
ORF01197	SAG1062	282	lipoate-protein ligase A family protein
ORF01198	SAG1063	228	flavoprotein-related protein
ORF01199	SAG1064	180	flavoprotein family protein
ORF01200	SAG1065	190	membrane protein, putative
ORF01201	SAG1066	572	phosphoglucomutase
ORF01202	SAG1067	178	IS861, transposase OrfA
ORF01203	SAG1068	277	IS861, transposase OrfB
ORF01204	SAG1069	65	hypothetical protein
ORF01205	SAG1070	577	ABC transporter, ATP-binding/permease protein
ORF01206	SAG1071	573	ABC transporter, ATP-binding/permease protein
ORF01207	SAG1072	200	conserved hypothetical protein
ORF01208	SAG1073	325	conserved hypothetical protein
ORF01209	SAG1074	418	Serine hydroxymethyltransferase
ORF01210	SAG1075	183	Sua5/YciO/YrdC/YwlC family protein
ORF01211	SAG1076	276	modification methylase, HemK family
ORF01212	SAG1077	359	peptide chain release factor 1
ORF01213	SAG1078	189	thymidine kinases
ORF01214	SAG1079	60	4-oxalocrotonate tautomerase
ORF01215	SAG1080	47	hypothetical protein
ORF01216	SAG1081	312	ApbE family protein
ORF01217	SAG1082	200	conserved hypothetical protein
ORF01218	SAG1083	411	conserved hypothetical protein
ORF01219	SAG1084	262	formate/nitrite transporter family protein
ORF01220	SAG1085	424	xanthine permease
ORF01221	SAG1086	193	xanthine phosphoribosyltransferase
ORF01222	SAG1087	327	guanosine monophosphate reductase
ORF01223	SAG1088	446	drug resistance transporter, EmrB/QacA family, putative

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01224	SAG1089	230	conserved hypothetical protein
ORF01225	SAG1090	666	potassium uptake protein, putative
ORF01226	SAG1091	216	oxidoreductase, short chain dehydrogenase/reductase family, FRAMESHIFT
ORF01227	SAG1092	330	phosphate acetyltransferase
ORF01228	SAG1093	294	ribosomal large subunit pseudouridine synthase, RluD subfamily
ORF01229	SAG1094	278	conserved hypothetical protein
ORF01230	SAG1095	223	GTP pyrophosphokinase family protein
ORF01231	SAG1096	190	conserved hypothetical protein
ORF01232	SAG1097	324	ribose-phosphate pyrophosphokinase
ORF01233	SAG1098	371	cysteine desulphurase
ORF01234	SAG1099	115	conserved hypothetical protein
ORF01235	SAG1100	210	DNA-binding protein
ORF01236	SAG1101	226	DNA repair protein RadC
ORF01237	SAG1102	377	membrane protein, putative
ORF01238	SAG1103	478	6-phospho-beta-glucosidase
ORF01239	SAG1104	204	platelet activating factor, putative
ORF01240	SAG1105	273	hydrolase, haloacid dehalogenase-like family
ORF01241	SAG1106	309	transcriptional regulator, AraC family, putative
ORF01242	SAG1107	510	voltage-gated chloride channel family protein
ORF01243	SAG1108	357	spermidine/putrescine ABC transporter, spermidine/putrescine-binding protein
ORF01244	SAG1109	258	spermidine/putrescine ABC transporter, permease protein
ORF01245	SAG1110	264	spermidine/putrescine ABC transporter, permease protein
ORF01246	SAG1111	384	spermidine/putrescine ABC transporter, ATP-binding protein
ORF01247	SAG1112	300	UDP-N-acetylenolpyruvoylglucosamine reductase
ORF01248	SAG1113	162	2-amino-4-hydroxy-6-hydroxymethyldihydropteridine pyrophosphokinase
ORF01249	SAG1114	120	dihydroneopterin aldolase
ORF01250	SAG1115	267	dihydropteroate synthase
ORF01251	SAG1116	187	GTP cyclohydrolase I
ORF01252	SAG1117	420	folylpolyglutamate synthase
ORF01253	SAG1118	295	rarD protein
ORF01254	SAG1119	288	homoserine kinase
ORF01255	SAG1120	427	homoserine dehydrogenase
ORF01256	SAG1121	295	polysaccharide deacetylase family protein
ORF01257	SAG1122	515	transporter, BCCT family protein
ORF01258	SAG1123	34	hypothetical protein
ORF01259	SAG1124	458	aldehyde dehydrogenase family protein
ORF01260	SAG1125	335	membrane protein
ORF01261	SAG1126	228	conserved hypothetical protein
ORF01262	SAG1127	113	conserved hypothetical protein, FRAMESHIFT
ORF01263		187	hypothetical protein
ORF01264	SAG1128	65	transcriptional regulator, Cro/CI family
ORF01265	SAG1129	36	hypothetical protein
ORF01266	SAG1130	49	hypothetical protein
ORF01268	SAG1131	164	thiol peroxidase
ORF01269	SAG1132	219	conserved hypothetical protein
ORF01272	SAG1133	254	conserved hypothetical protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01273	SAG1134	213	transcriptional regulator, GntR family/potassioum uptake protein, TrkA family
ORF01274	SAG1135	183	gls24 protein, putative
ORF01275	SAG1136		conserved hypothetical protein FRAMESHIFT
ORF01276	SAG1137	180	gls24 protein, putative
ORF01277	SAG1138	64	conserved hypothetical protein
ORF01279	SAG1139	193	conserved hypothetical protein
ORF01280	SAG1140	82	conserved hypothetical protein
ORF01281	SAG1141	112	conserved hypothetical protein
ORF01282	SAG1142	759	ATP-dependent DNA helicase PcrA
ORF01283	SAG1143	100	conserved hypothetical protein, FRAMESHIFT
ORF01284	SAG1144	441	uracil permease
ORF01285	SAG1145	448	sodium:alanine symporter family protein
ORF01286	SAG1146	411	cation efflux family protein
ORF01287	SAG1147	130	conserved hypothetical protein
ORF01288	SAG1148	231	membrane protein, putative
ORF01289	SAG1149	207	conserved hypothetical protein
ORF01290	SAG1150	400	ribosomal protein S1
ORF01291	SAG1151	76	conserved hypothetical protein
ORF01292	SAG1152	340	branched-chain amino acid aminotransferase
ORF01294	SAG1153	819	DNA topoisomerase IV, A subunit
ORF01295	SAG1154	653	DNA topoisomerase IV, B subunit
ORF01296	SAG1155	207	conserved hypothetical protein TIGR00023
ORF01297	SAG1156	217	uracil-DNA glycosylase
ORF01298	SAG1157	161	conserved hypothetical protein
ORF01299	SAG1158	413	CMP-N-acetylneuraminic acid synthetase NeuA
ORF01300	SAG1159	209	neuD protein
ORF01301	SAG1160	384	UDP-N-acetylglucosamine-2-epimerase NeuC
ORF01302	SAG1161	341	N-acetyl neuramic acid synthetase NeuB
ORF01303	SAG1162	466	cpsL protein
ORF01304	SAG1163	318	cpsVK protein
ORF01305	SAG1164	321	cpsVJ protein
ORF01306	SAG1165	327	cpsVO protein
ORF01307	SAG1166	295	cpsVN protein
ORF01308	SAG1167	241	cpsVM protein
ORF01309	SAG1168	364	cpsVH protein
ORF01310	SAG1169	163	CpsVG
ORF01311	SAG1170	149	CpsF
ORF01312	SAG1171	462	CpsE
ORF01313	SAG1172	229	CpsD protein
ORF01314	SAG1173	230	cpsC protein
ORF01315	SAG1174	243	capsular polysaccharide biosynthesis protein CpsB
ORF01316	SAG1175	485	capsular polysaccharide biosynthesis protein CpsA
ORF01317	SAG1176	290	capsular polysaccharide synthesis operon transcriptional regulator CpsY
ORF01318	SAG1177	255	cpsIaS protein
ORF01319	SAG1178	236	purine nucleoside phosphorylase
ORF01320	SAG1179	418	voltage-gated chloride channel family protein, putative
ORF01321	SAG1180	269	purine nucleoside phosphorylase
ORF01322	SAG1181	135	arsenate reductase
ORF01323	SAG1182	403	phosphopentomutase
ORF01324	SAG1183	223	ribose 5-phosphate isomerase

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01326	SAG1184	236	conserved hypothetical protein
ORF01327	SAG1185	262	tributyrin esterase
ORF01328	SAG1186	553	metallo-beta-lactamase superfamily protein
ORF01329	SAG1187	253	ABC transporter, ATP-binding protein
ORF01330	SAG1188	287	ABC transporter, permease protein
ORF01331	SAG1189	334	conserved hypothetical protein
ORF01332	SAG1190	551	adherence and virulence protein A
ORF01333	SAG1191	239	alpha-acetolactate decarboxylase
ORF01334	SAG1192	560	acetolactate synthase, catabolic
ORF01335	SAG1193	408	TPR domain protein
ORF01336	SAG1194	396	membrane protein
ORF01337	SAG1195	153	MutT/nudix family protein
ORF01338	SAG1196	160	mutator MutT protein
ORF01339	SAG1197	1072	hyaluronidase
ORF01340	SAG1198	348	dTDP-glucose 4,6-dehydratase
ORF01341	SAG1199	197	dTDP-4-dehydrorhamnose 3,5-epimerase
ORF01342	SAG1200	289	glucose-1-phosphate thymidylyltransferase
ORF01343	SAG1201	367	iminodiacetate oxidase, putative
ORF01344	SAG1202	262	conserved hypothetical protein TIGR00486
ORF01345	SAG1203	227	conserved hypothetical protein
ORF01346	SAG1204	226	DNA replication protein Dnad, putative
ORF01347	SAG1205	172	adenine phosphoribosyltransferase
ORF01348	SAG1206	854	conserved domain protein
ORF01349	SAG1207	32	hypothetical protein
ORF01350	SAG1208	732	single-stranded-DNA-specific exonuclease RecJ
ORF01351	SAG1209	253	oxidoreductase, short chain dehydrogenase/reductase family
ORF01352	SAG1210	309	metallo-beta-lactamase superfamily protein
ORF01353	SAG1211	215	conserved hypothetical protein
ORF01354	SAG1212	412	GTP-binding protein HflX
ORF01355	SAG1213	296	tRNA delta(2)-isopentenylpyrophosphate transferase
ORF01356	SAG1214	58	hypothetical protein
ORF01357	SAG1215	305	exfoliative toxin A, putative
ORF01358	SAG1216	1252	pullulanase, putative
ORF01361	SAG1217		conserved hypothetical protein, FRAMESHIFT
ORF01362	SAG1218	194	conserved hypothetical protein
ORF01363	SAG1219	468	peptidase, M20/M25/M40 family
ORF01364	SAG1220	200	nitroreductase family protein
ORF01365	SAG1221		glycerophosphoryl diester phosphodiesterase, putative, POINT MUTATION
ORF01367	SAG1222	593	excinuclease ABC, C subunit
ORF01368	SAG1223	255	conserved hypothetical protein
ORF01369	SAG1224	446	MATE efflux family protein
ORF01370	SAG1225	136	conserved hypothetical protein
ORF01371	SAG1226	165	conserved hypothetical protein
ORF01372	SAG1227	198	conserved hypothetical protein
ORF01373	SAG1228	96	ISSdy1, transposase OrfA
ORF01374	SAG1229	259	ISSdy1, transposase OrfB
ORF01375	SAG1230	96	conserved hypothetical protein
ORF01377	SAG1231		transposase OrfB, IS3 family, degenerate FRAMESHIFT
ORF01379	SAG1232	77	transposase OrfB, IS3 family, truncation
ORF01380	SAG1233	822	streptococcal histidine triad family protein
ORF01381	SAG1234	306	laminin-binding surface protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01382	SAG1235	425	GBSi1, group II intron, maturase
ORF01383	SAG1236		c5a peptidase precursor FRAMESHIFT
ORF01384	SAG1237	444	hypothetical protein
ORF01385	SAG1238	202	hypothetical protein
ORF01386	SAG1239	76	conserved hypothetical protein
ORF01387	SAG1240	125	conserved hypothetical protein, truncation
ORF01388	SAG1241	78	transposase OrfA, IS3 family
ORF01389	SAG1242	67	transposase OrfB, IS3 family, truncation
ORF01390	SAG1243	96	ISSdy1, transposase OrfA FRAMESHIFT
ORF01391	SAG1244	259	ISSdy1, transposase OrfB
ORF01392	SAG1245	38	hypothetical protein
ORF01393	SAG1246	389	hypothetical protein
ORF01394	SAG1247	399	integrase, phage family
ORF01395	SAG1248	75	conserved hypothetical protein
ORF01396	SAG1249	74	transcriptional regulator, Cro/CI family
ORF01397	SAG1250	621	Tn5252, relaxase
ORF01398	SAG1251	121	Tn5252, Orf 9 protein
ORF01399	SAG1252	120	Tn5252, Orf 10 protein
ORF01401	SAG1253	435	transposase, ISL3 family
ORF01403	SAG1254	546	mercuric reductase
ORF01404	SAG1255	130	mercuric resistance operon regulatory protein MerR
ORF01406	SAG1256	142	IS861, transposase OrfB, truncation
ORF01407	SAG1257	709	cation-transporting ATPase, E1-E2 family
ORF01408	SAG1258	122	cadmium efflux system accessory protein
ORF01409	SAG1259	99	conserved hypothetical protein
ORF01410	SAG1260	262	hypothetical protein
ORF01411	SAG1261	198	conserved hypothetical protein
ORF01412	SAG1262	695	cation-transporting ATPase, E1-E2 family
ORF01414	SAG1263		conserved domain protein, FRAMESHIFT
ORF01415	SAG1264	148	transcriptional repressor CopY, putative
ORF01416	SAG1265	206	cadmium resistance transporter, putative
ORF01417	SAG1266	152	hypothetical protein
ORF01418	SAG1267	108	hypothetical protein
ORF01419	SAG1268	230	repressor protein, putative
ORF01420	SAG1269	44	hypothetical protein
ORF01421	SAG1270	471	ImpB/MucB/SamB family protein
ORF01423	SAG1271	116	conserved hypothetical protein
ORF01424	SAG1272	102	conserved hypothetical protein
ORF01425	SAG1273	118	conserved hypothetical protein
ORF01426	SAG1274	129	conserved hypothetical protein
ORF01427	SAG1275	75	hypothetical protein
ORF01428	SAG1276	358	conserved hypothetical protein
ORF01430	SAG1277	163	hypothetical protein
ORF01431	SAG1278	96	hypothetical protein
ORF01432	SAG1279	99	conserved domain protein
ORF01433	SAG1280	2274	Helicases conserved C-terminal domain protein
ORF01434	SAG1281	183	hypothetical protein
ORF01435	SAG1282	63	lipoprotein, putative
ORF01436	SAG1283	1631	cell wall surface anchor family protein
ORF01437	SAG1284	196	abortive infection protein AbiGI
ORF01438	SAG1285	281	abortive infection protein AbiGII
ORF01439	SAG1286	933	conserved hypothetical protein
ORF01440	SAG1287	776	conserved hypothetical protein
ORF01441	SAG1288	117	conserved hypothetical protein, DEGENERATE

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01442	SAG1289	284	conserved hypothetical protein
ORF01443	SAG1290	80	hypothetical protein
ORF01444	SAG1291	605	Tn5252, Orf 21 protein, internal deletion
ORF01445	SAG1292	162	hypothetical protein
ORF01446	SAG1293	194	protease, putative
ORF01447	SAG1294	77	conserved hypothetical protein
ORF01449	SAG1295	127	conserved hypothetical protein
ORF01450	SAG1296	142	conserved hypothetical protein
ORF01451	SAG1297	451	type II DNA modification methyltransferase Spn5252IP
ORF01452	SAG1298	31	hypothetical protein
ORF01453	SAG1299	272	conserved hypothetical protein
ORF01454	SAG1300	57	conserved hypothetical protein
ORF01455	SAG1301	121	ribosomal protein L7/L12
ORF01456	SAG1302	166	ribosomal protein L10
ORF01458	SAG1303	702	ATP-dependent Clp protease, ATP-binding subunit
ORF01459	SAG1304	32	hypothetical protein
ORF01460	SAG1305	314	homocysteine S-methyltransferase MmuM, putative
ORF01461	SAG1306	458	amino acid permease
ORF01463	SAG1307	216	hypothetical protein
ORF01464	SAG1308	167	hypothetical protein
ORF01465	SAG1309	30	hypothetical protein
ORF01466	SAG1310	182	transcriptional regulator, TetR family
ORF01467	SAG1311	198	GTP-binding protein
ORF01468	SAG1312	408	ATP-dependent Clp protease, ATP-binding subunit ClpX
ORF01469	SAG1313	56	conserved hypothetical protein
ORF01470	SAG1314	164	dihydrofolate reductase
ORF01471	SAG1315	279	thymidylate synthase
ORF01472	SAG1316	390	HMG-CoA synthase
ORF01473	SAG1317	427	3-hydroxy-3-methylglutaryl-CoA reductase
ORF01474	SAG1318	149	conserved hypothetical protein
ORF01475	SAG1319	187	hemolysin III, putative
ORF01476	SAG1320	304	conserved hypothetical protein TIGR00147
ORF01477	SAG1321	284	glutathione S-transferase family protein
ORF01478	SAG1322	72	conserved domain protein
ORF01479	SAG1323	331	isopentenyl-diphosphate delta-isomerase
ORF01480	SAG1324	330	phosphomevalonate kinase
ORF01481	SAG1325	314	diphosphomevalonate decarboxylase
ORF01482	SAG1326	292	mevalonate kinase, putative
ORF01483	SAG1327	409	sensor histidine kinase
ORF01484	SAG1328	228	DNA-binding response regulator
ORF01485	SAG1329	208	GTP pyrophosphokinase family protein
ORF01486	SAG1330	68	hypothetical protein
ORF01488	SAG1331	979	R5 protein
ORF01489	SAG1332	146	transcriptional regulator, MarR family, putative
ORF01490	SAG1333	690	5'-nucleotidase family protein
ORF01491	SAG1334	136	polypeptide deformylase, putative
ORF01492	SAG1335	449	NADP-specific glutamate dehydrogenase
ORF01494	SAG1336	169	conserved hypothetical protein
ORF01495	SAG1337	589	ABC transporter, ATP-binding/permease protein
ORF01496	SAG1338	579	ABC transporter, ATP-binding/permease protein
ORF01497	SAG1339	157	acetyltransferase, GNAT family

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01498	SAG1340	622	ABC transporter, ATP-binding protein
ORF01499	SAG1341	402	polyA polymerase family protein
ORF01500	SAG1342	282	DegV family protein
ORF01501	SAG1343	126	conserved hypothetical protein
ORF01502	SAG1344	177	hypothetical protein
ORF01503	SAG1345	164	conserved hypothetical protein
ORF01504	SAG1346	641	PTS system, fructose specific IIABC components
ORF01505	SAG1347	303	1-phosphofructokinase
ORF01506	SAG1348	247	lactose phosphotransferase system repressor
ORF01507	SAG1349	411	beta-lactam resistance factor
ORF01508	SAG1350	544	surface antigen-related protein
ORF01509	SAG1351	307	2-dehydropantoate 2-reductase, putative
ORF01510	SAG1352	356	regulatory protein, putative
ORF01511	SAG1353	330	pyridine nucleotide-disulphide oxidoreductase family protein
ORF01512	SAG1354	251	tRNA (guanine-N1)-methyltransferase
ORF01513	SAG1355	172	16S rRNA processing protein RimM
ORF01515	SAG1356	503	transcriptional regulator, RofA family
ORF01516	SAG1357	80	KH domain protein
ORF01517	SAG1358	90	ribosomal protein S16
ORF01518	SAG1359	415	permease, putative
ORF01519	SAG1360	236	ABC transporter, ATP-binding protein
ORF01520	SAG1361	414	conserved hypothetical protein
ORF01522	SAG1362	532	carbamoyl-phosphate synthase, large subunit, putative
ORF01523	SAG1363	356	carbamoyl-phosphate synthase, small subunit
ORF01524	SAG1364	173	pyrimidine operon regulatory protein
ORF01525	SAG1365	296	ribosomal large subunit pseudouridine synthase, RluD subfamily
ORF01526	SAG1366	154	lipoprotein signal peptidase
ORF01527	SAG1367	301	transcriptional regulator, LysR family
ORF01528	SAG1368	94	ribosomal protein L27
ORF01529	SAG1369	112	conserved hypothetical protein
ORF01530	SAG1370	104	ribosomal protein L21
ORF01531	SAG1371	392	conserved hypothetical protein
ORF01532	SAG1372	404	thiamine biosynthesis protein ThiI
ORF01533	SAG1373	381	cysteine desulphurase
ORF01535	SAG1374	150	conserved hypothetical protein
ORF01536	SAG1375	449	glutathione reductase
ORF01537	SAG1376	111	conserved hypothetical protein
ORF01538	SAG1377	388	chorismate synthase
ORF01539	SAG1378	355	3-dehydroquinate synthase
ORF01540	SAG1379	225	3-dehydroquinate dehydratase
ORF01541	SAG1380	385	conserved hypothetical protein
ORF01542	SAG1381	714	sulfatase
ORF01543	SAG1382	119	ribosomal protein L20
ORF01544	SAG1383	66	ribosomal protein L35
ORF01545	SAG1384	176	translation initiation factor IF-3
ORF01546	SAG1385	227	cytidylate kinase
ORF01547	SAG1386	174	conserved hypothetical protein
ORF01548	SAG1387	65	ferredoxin, 4Fe-4S
ORF01549	SAG1388	163	conserved hypothetical protein
ORF01550	SAG1389	406	peptidase t
ORF01551	SAG1390	544	polysaccharide biosynthesis protein, putative

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01552	SAG1391	484	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase
ORF01553	SAG1392	264	iron compound ABC transporter, ATP-binding protein
ORF01554	SAG1393	310	iron compound ABC transporter, substrate-binding protein
ORF01555	SAG1394	341	iron compound ABC transporter, permease protein
ORF01556	SAG1395	333	iron compound ABC transporter, permease protein
ORF01557	SAG1396	217	conserved hypothetical protein
ORF01558	SAG1397	311	inorganic pyrophosphatase, manganese-dependent
ORF01559	SAG1398	262	pyruvate formate-lyase-activating enzyme
ORF01560	SAG1399	444	CBS domain protein
ORF01561	SAG1400	188	conserved hypothetical protein
ORF01563	SAG1401	311	conserved hypothetical protein TIGR01212
ORF01564	SAG1402	213	PAP2 family protein
ORF01565	SAG1403	194	membrane protein, putative
ORF01566	SAG1404	308	cell wall surface anchor family protein
ORF01567	SAG1405	294	sortase family protein
ORF01568	SAG1406	293	sortase family protein
ORF01569	SAG1407	705	cell wall surface anchor family protein
ORF01570	SAG1408	901	cell wall surface anchor family protein
ORF01571	SAG1409	326	transcriptional regulator, RofA family FRAMESHIFT
ORF01572	SAG1410	379	glycosyl transferase, group 1 family protein
ORF01573	SAG1411	282	exopolysaccharide biosynthesis protein, putative
ORF01574	SAG1412	474	exopolysaccharide biosynthesis protein, putative
ORF01575	SAG1413	454	hypothetical protein
ORF01576	SAG1414	308	glycosyl transferase, group 2 family protein
ORF01577	SAG1415	311	glycosyl transferase, group 2 family protein
ORF01578	SAG1416	352	dTDP-glucose 4,6-dehydratase, putative
ORF01579	SAG1417	240	4-diphosphocytidyl-2C-methyl-D-erythritol synthase, putative
ORF01580	SAG1418	259	licD protein, putative
ORF01581	SAG1419	577	hypothetical protein
ORF01582	SAG1420	117	conserved hypothetical protein
ORF01583	SAG1421	243	glycosyl transferase, group 2 family protein
ORF01584	SAG1422	313	glycosyl transferase, group 2 family protein
ORF01585	SAG1423	384	conserved hypothetical protein
ORF01586	SAG1424	284	dTDP-4-dehydrorhamnose reductase
ORF01587	SAG1425	113	conserved hypothetical protein
ORF01589	SAG1426	369	RNA polymerase sigma-70 factor
ORF01590	SAG1427	602	DNA primase
ORF01591	SAG1428	125	large conductance mechanosensitive channel protein
ORF01592	SAG1429	58	ribosomal protein S21
ORF01593	SAG1430	167	conserved hypothetical protein
ORF01594	SAG1431	268	amino acid ABC transporter, amino acid-binding protein
ORF01596	SAG1432	347	ammonium transporter family protein
ORF01597	SAG1433	375	conserved hypothetical protein
ORF01598	SAG1434	328	rhodanese family protein
ORF01599	SAG1435	101	conserved hypothetical protein
ORF01600	SAG1436	457	glycerol-3-phosphate transporter, putative

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01601	SAG1437	55	hypothetical protein
ORF01602	SAG1438	754	glycogen phosphorylase
ORF01603	SAG1439	498	4-alpha-glucanotransferase
ORF01604	SAG1440	342	maltose operon repressor MalR, putative
ORF01605	SAG1441	415	maltose/maltodextrin ABC transporter, maltose/maltodextrin-binding protein
ORF01606	SAG1442	456	maltose ABC transporter, permease protein
ORF01607	SAG1443	278	maltose ABC transporter, permease protein
ORF01608	SAG1444	490	proton/peptide symporter family protein
ORF01610	SAG1445		MutT/nudix family protein, FRAMESHIFT
ORF01611	SAG1446	62	hypothetical protein
ORF01612	SAG1447	441	conserved hypothetical protein
ORF01613	SAG1448	502	glycosyl transferase, group 1 family protein
ORF01614	SAG1449	795	preprotein translocase SecA subunit, putative
ORF01615	SAG1450	330	conserved domain protein
ORF01617	SAG1451	494	conserved hypothetical protein
ORF01618	SAG1452	514	conserved hypothetical protein
ORF01619	SAG1453	409	preprotein translocase SecY family protein
ORF01621	SAG1454	398	conserved hypothetical protein
ORF01622	SAG1455	295	glycosyl transferase, group 2 family protein
ORF01623	SAG1456	312	glycosyl transferase, family 8, degenerate
ORF01624	SAG1457	129	IS1381, transposase OrfB
ORF01625	SAG1458	127	IS1381, transposase OrfA
ORF01626	SAG1459	413	glycosyl transferase family 8
ORF01627	SAG1460	401	glycosyl transferase, family 8
ORF01628	SAG1461	335	conserved hypothetical protein
ORF01630	SAG1462	970	cell wall surface anchor family protein
ORF01632	SAG1463		transcriptional regulator, RofA family POINT MUTATION
ORF01634	SAG1464	663	excinuclease ABC, B subunit
ORF01635	SAG1465	306	protease, putative
ORF01636	SAG1466	727	glutamine ABC transporter, glutamine-binding protein/permease protein, putative
ORF01637	SAG1467	246	glutamine ABC transporter, ATP-binding protein, GlnQ putative
ORF01638	SAG1468	116	conserved hypothetical protein
ORF01639	SAG1469	52	conserved hypothetical protein
ORF01640	SAG1470	437	GTP-binding protein, GTP1/Obg family
ORF01641	SAG1471	42	conserved hypothetical protein
ORF01643	SAG1472	413	aminopeptidase PepS
ORF01645	SAG1473	192	cell wall surface anchor family protein
ORF01646	SAG1474	680	amidase family protein
ORF01647	SAG1475	240	ribosomal small subunit pseudouridine synthase A
ORF01648	SAG1476	280	oxidoreductase, aldo/keto reductase family
ORF01650	SAG1477	224	nitroreductase family protein
ORF01651	SAG1478	130	lactoylglutathione lyase
ORF01652	SAG1479	308	glycosyl transferase, group 2 family protein
ORF01653	SAG1480	462	amino acid permease
ORF01654	SAG1481	155	SsrA-binding protein
ORF01655	SAG1482	801	exoribonuclease, VacB/Rnb family
ORF01657	SAG1483	78	preprotein translocase, SecG subunit
ORF01658	SAG1485	389	multi-drug resistance protein
ORF01660	SAG1486	548	hypothetical protein
ORF01661	SAG1487	233	ABC transporter, ATP binding protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01662	SAG1488	195	dephospho-CoA kinase
ORF01663	SAG1489	273	formamidopyrimidine-DNA glycosylase
ORF01665	SAG1490	282	transcriptional regulator, MutR family
ORF01666	SAG1491	530	hypothetical protein
ORF01667	SAG1492	58	hypothetical protein
ORF01668	SAG1493		hypothetical protein
ORF01670	SAG1494	32	hypothetical protein
ORF01672	SAG1495	81	protease, putative, POINT MUTATION
ORF01673	SAG1496	110	hypothetical protein
ORF01674	SAG1497	37	hypothetical protein
ORF01675	SAG1498	133	hypothetical protein
ORF01677	SAG1499	299	GTP-binding protein Era
ORF01678	SAG1500	132	diacylglycerol kinase
ORF01679	SAG1501	161	conserved hypothetical protein TIGR00043
ORF01680	SAG1502	268	tetracenomycin polyketide synthesis O-methyltransferase TcmP, putative
ORF01681	SAG1503	39	hypothetical protein
ORF01682	SAG1504	38	hypothetical protein
ORF01683	SAG1505	158	MutT/nudix family protein
ORF01684	SAG1506	267	hypothetical protein
ORF01685	SAG1507	345	PhoH family protein
ORF01686	SAG1508	590	67 kDa Myosin-crossreactive streptococcal antigen
ORF01687	SAG1509	71	conserved hypothetical protein
ORF01688	SAG1510	169	peptide methionine sulfoxide reductase
ORF01689	SAG1511	284	conserved hypothetical protein
ORF01690	SAG1512	185	ribosome recycling factor
ORF01691	SAG1513	242	uridylate kinase
ORF01692	SAG1514	226	peptide ABC transporter, ATP-binding protein
ORF01693	SAG1515	262	peptide ABC transporter, ATP-binding protein
ORF01694	SAG1516	255	peptide ABC transporter, permease protein
ORF01695	SAG1517	314	peptide ABC transporter, permease protein
ORF01696	SAG1518	525	peptide ABC transporter, peptide-binding protein
ORF01697	SAG1519	229	ribosomal protein L1
ORF01698	SAG1520	141	ribosomal protein L11
ORF01699	SAG1521	388	transposase, IS30 family, putative
ORF01700	SAG1522	460	transporter, major facilitator family
ORF01702	SAG1523	404	peptidase, M20/M25/M40 family
ORF01703	SAG1524	294	transcriptional regulator, LysR family
ORF01704	SAG1525	117	conserved hypothetical protein
ORF01705	SAG1526	178	IS861, transposase OrfA
ORF01706	SAG1527	277	IS861, transposase OrfB
ORF01707	SAG1528	571	chorismate binding enzyme
ORF01708	SAG1529	785	FtsK/SpoIIIE family protein
ORF01709	SAG1530	267	peptidyl-prolyl cis-trans isomerase, cyclophilin-type
ORF01710	SAG1531	277	manganese ABC transporter, permease protein
ORF01711	SAG1532	238	manganese ABC transporter, ATP-binding protein
ORF01712	SAG1533	308	manganese ABC transporter, manganese-binding adhesion liprotein
ORF01713	SAG1534	215	iron-dependent transcriptional regulator
ORF01714	SAG1535	229	5-methylthioadenosine nucleosidase/S-adenosylhomocysteine nucleosidase
ORF01715	SAG1536	89	conserved hypothetical protein
ORF01716	SAG1537	184	MutT/nudix family protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01718	SAG1538	459	UDP-N-acetylglucosamine pyrophosphorylase
ORF01719	SAG1539	31	hypothetical protein
ORF01720	SAG1540	137	conserved hypothetical protein
ORF01721	SAG1541	125	glyoxalase family protein
ORF01722	SAG1542	318	oxidoreductase, Gfo/Idh/MocA family
ORF01724	SAG1543		conserved hypothetical protein, FRAMESHIFT
ORF01725	SAG1544	232	gluconate 5-dehydrogenase, putative
ORF01726	SAG1545	78	conserved hypothetical protein
ORF01727	SAG1546	82	conserved hypothetical protein
ORF01729	SAG1547	166	acetyltransferase, GNAT family
ORF01730	SAG1548	422	glycosyl transferase, group 2 family protein
ORF01731	SAG1549	127	IS1381, transposase OrfA
ORF01732	SAG1550	129	IS1381, transposase OrfB
ORF01733	SAG1551	67	hypothetical protein
ORF01734	SAG1552	719	conserved hypothetical protein
ORF01735	SAG1553	477	hypothetical protein
ORF01736	SAG1554	225	hypothetical protein
ORF01737	SAG1555	231	hypothetical protein
ORF01738	SAG1556	445	branched-chain amino acid transport system II carrier protein
ORF01739	SAG1557	665	methionyl-tRNA synthetase
ORF01740	SAG1558	291	tellurite resistance protein TehB
ORF01741	SAG1559	231	membrane protein, putative
ORF01742	SAG1560	40	hypothetical protein
ORF01743	SAG1561	405	PTS system component, putative
ORF01744	SAG1562	280	conserved hypothetical protein
ORF01745	SAG1563	275	exodeoxyribonuclease
ORF01746	SAG1564	118	conserved hypothetical protein
ORF01747	SAG1565	158	methylated-DNA--protein-cysteine S-methyltransferase
ORF01748	SAG1566	393	D-isomer specific 2-hydroxyacid dehydrogenase family protein
ORF01749	SAG1567	182	acetyltransferase, GNAT family
ORF01750	SAG1568		phosphoserine aminotransferase FRAMESHIFT
ORF01752	SAG1569	211	copper homeostasis protein CutC, putative
ORF01753	SAG1570	34	conserved hypothetical protein
ORF01754	SAG1571	53	hypothetical protein
ORF01755	SAG1572	287	tetrapyrrole methylase family protein
ORF01756	SAG1573	108	conserved hypothetical protein
ORF01758	SAG1574	287	DNA polymerase III, delta prime subunit, putative
ORF01759	SAG1575	211	thymidylate kinase
ORF01761	SAG1576	267	transposase, IS30 family, putative, truncation
ORF01763	SAG1577	219	AcuB family protein
ORF01764	SAG1578	236	branched-chain amino acid ABC transporter, ATP-binding protein
ORF01765	SAG1579	254	branched-chain amino acid ABC transporter, ATP-binding protein
ORF01766	SAG1580	317	branched-chain amino acid ABC transporter, permease protein
ORF01767	SAG1581	289	branched-chain amino acid ABC transporter, permease protein
ORF01769	SAG1582	388	branched-chain amino acid ABC transporter, amino acid-binding protein
ORF01770	SAG1583	81	conserved hypothetical protein
ORF01772	SAG1584	377	IS1548, transposase

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01773	SAG1585	196	ATP-dependent Clp protease, proteolytic subunit ClpP
ORF01774	SAG1586	209	uracil phosphoribosyltransferase
ORF01775	SAG1587	389	aminotransferase, class I
ORF01777	SAG1588	182	RNA methyltransferase, TrmH family, group 2
ORF01778	SAG1589	450	amino acid permease, putative
ORF01779	SAG1590	449	potassium uptake protein, Trk family
ORF01780	SAG1591	475	cation uptake protein, Trk family
ORF01781	SAG1592	83	conserved hypothetical protein TIGR00278
ORF01782	SAG1593	240	ribosomal large subunit pseudouridine synthase B
ORF01783	SAG1594	194	conserved hypothetical protein TIGR00281
ORF01784	SAG1595	235	Uncharacterized ACR, COG1354
ORF01785	SAG1596	246	integrase/recombinase, phage integrase family
ORF01786	SAG1597	157	CBS domain protein
ORF01787	SAG1598	173	conserved hypothetical protein
ORF01788	SAG1599	324	HAM1 protein
ORF01789	SAG1600	264	glutamate racemase
ORF01790	SAG1601	79	conserved hypothetical protein
ORF01791	SAG1602	180	membrane protein, putative
ORF01792	SAG1603	173	transcriptional regulator, biotin repressor family
ORF01793	SAG1604	229	membrane protein, putative
ORF01794	SAG1605	167	conserved hypothetical protein
ORF01795	SAG1606	247	RNA methyltransferase, TrmH family
ORF01796	SAG1607	92	acylphosphatase
ORF01797	SAG1608	310	membrane protein, putative
ORF01799	SAG1609	221	amino acid ABC transporter, permease protein
ORF01800	SAG1610	285	amino acid ABC transporter, substrate-binding protein
ORF01801	SAG1611	486	amidase family protein
ORF01802	SAG1612	160	transcription elongation factor GreA
ORF01803	SAG1613	600	Uncharacterized BCR, YceG family COG1559, putative
ORF01804	SAG1614	167	acetyltransferase, GNAT family
ORF01805	SAG1615	443	UDP-N-acetylmuramate--alanine ligase
ORF01806	SAG1616	205	conserved hypothetical protein
ORF01807	SAG1617	32	hypothetical protein
ORF01808	SAG1618	1032	Snf2 family protein
ORF01810	SAG1619	377	IS1548, transposase
ORF01811	SAG1620	436	phosphoglycerate dehydrogenase-related protein
ORF01812	SAG1621	300	primosomal protein DnaI
ORF01813	SAG1622	391	conserved hypothetical protein
ORF01814	SAG1623	159	conserved hypothetical protein TIGR00244
ORF01815	SAG1624	501	sensor histidine kinase CsrS
ORF01816	SAG1625	229	DNA-binding response regulator CsrR
ORF01817	SAG1626	177	conserved hypothetical protein
ORF01818	SAG1627	296	heat shock protein HtpX
ORF01820	SAG1628	184	lemA protein
ORF01821	SAG1629	237	glucose-inhibited division protein B
ORF01822	SAG1630	459	sodium transport family protein
ORF01823	SAG1631	223	potassium uptake protein, Trk family, putative
ORF01824	SAG1632	276	cobalt transport family protein
ORF01825	SAG1633	558	ABC transporter, ATP-binding protein
ORF01826	SAG1634	212	conserved hypothetical protein
ORF01827	SAG1635	402	sodium:dicarboxylate symporter family protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01828	SAG1636	455	branched-chain amino acid transport system II carrier protein
ORF01829	SAG1637	351	alcohol dehydrogenase, zinc-containing
ORF01830	SAG1638	230	ABC transporter, permease protein
ORF01831	SAG1639	356	ABC transporter, ATP-binding protein
ORF01832	SAG1640	458	peptidase, M20/M25/M40 family
ORF01833	SAG1641	274	lipoprotein,putative
ORF01834	SAG1642	277	ABC transporter, substrate-binding protein
ORF01835	SAG1643	229	glutamine amidotransferase, class I
ORF01836	SAG1644	37	hypothetical protein
ORF01837	SAG1645	238	conserved hypothetical protein TIGR01033
ORF01838	SAG1646	32	hypothetical protein
ORF01839	SAG1647	328	dihydroxyacetone kinase family protein
ORF01840	SAG1648	178	transcriptional regulator, TetR family, putative
ORF01842	SAG1649	37	hypothetical protein
ORF01843	SAG1650	329	dihydroxyacetone kinase family protein
ORF01844	SAG1651	192	dihydroxyacetone kinase family protein
ORF01845	SAG1652	124	conserved hypothetical protein
ORF01846	SAG1653	237	glycerol uptake facilitator protein
ORF01847	SAG1654	134	conserved hypothetical protein
ORF01848	SAG1655	237	transcriptional regulator, MerR family
ORF01849	SAG1656	369	conserved hypothetical protein
ORF01850	SAG1657	83	hypothetical protein
ORF01851	SAG1658	244	conserved hypothetical protein
ORF01852	SAG1659	118	iojap-related protein
ORF01853	SAG1660	173	isochorismatase family protein
ORF01854	SAG1661	195	conserved hypothetical protein TIGR00488
ORF01855	SAG1662	210	conserved hypothetical protein TIGR00482
ORF01856	SAG1663	105	conserved hypothetical protein TIGR00253
ORF01857	SAG1664	372	GTP-binding protein
ORF01858	SAG1665	177	hydrolase, haloacid dehalogenase-like family
ORF01859	SAG1666	295	membrane protein
ORF01860	SAG1667	480	glutamyl-tRNA(Gln) amidotransferase, B subunit
ORF01861	SAG1668	488	glutamyl-tRNA(Gln) amidotransferase, A subunit
ORF01862	SAG1669	100	glutamyl-tRNA(Gln) amidotransferase, C subunit
ORF01863	SAG1670	881	pyruvate phosphate dikinase
ORF01864	SAG1671	276	conserved hypothetical protein
ORF01865	SAG1672	170	CBS domain protein
ORF01866	SAG1673	377	3-hydroxyacyl-CoA dehydrogenase family protein
ORF01867	SAG1674	182	isochorismatase family protein
ORF01869	SAG1675	261	transcriptional regulator CodY, putative
ORF01870	SAG1676	403	aminotransferase, class I
ORF01871	SAG1677	137	universal stress protein family FRAMESHIFT
ORF01872	SAG1678	460	hydrolase, haloacid dehalogenase-like family
ORF01873	SAG1679	320	asparaginase family protein
ORF01874	SAG1680	292	shikimate 5-dehydrogenase
ORF01875	SAG1681	304	oxidoreductase, aldo/keto reductase family
ORF01876	SAG1682	671	ATP-dependent DNA helicase RecG
ORF01877	SAG1683	512	immunogenic secreted protein, putative
ORF01878	SAG1684	366	alanine racemase
ORF01879	SAG1685	119	holo-(acyl-carrier-protein) synthase
ORF01880	SAG1686	335	phospho-2-dehydro-3-deoxyheptonate aldolase
ORF01881	SAG1687	842	preprotein translocase, SecA subunit
ORF01882	SAG1688	315	mannose-6-phosphate isomerase, class I
ORF01883	SAG1689	293	fructokinase

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01885	SAG1690	639	PTS system IIABC components
ORF01886	SAG1691	479	sucrose-6-phosphate hydrolase
ORF01887	SAG1692	320	sucrose operon repressor ScrR
ORF01888	SAG1693	144	N utilization substance protein B
ORF01889	SAG1694	129	conserved hypothetical protein
ORF01890	SAG1695	186	translation elongation factor P
ORF01892	SAG1696	38	hypothetical protein
ORF01893	SAG1697	48	hypothetical protein
ORF01894	SAG1698	99	conserved hypothetical protein
ORF01895	SAG1699	30	hypothetical protein
ORF01896	SAG1700	76	hypothetical protein
ORF01897	SAG1701	56	hypothetical protein
ORF01898	SAG1702	41	hypothetical protein
ORF01899	SAG1703	54	hypothetical protein
ORF01900	SAG1704	150	cytidine/deoxycytidylate deaminase family protein
ORF01902	SAG1705		peptidase, M24 family POINT MUTATION
ORF01903	SAG1706	238	conserved hypothetical protein
ORF01904	SAG1707	499	drug resistance transporter, EmrB/QacA family
ORF01905	SAG1708	38	hypothetical protein
ORF01906	SAG1709	942	excinuclease ABC, A subunit
ORF01907	SAG1710	223	conserved hypothetical protein
ORF01908	SAG1711	314	magnesium transporter, CorA family
ORF01909	SAG1712	79	ribosomal protein S18
ORF01910	SAG1713	163	single-strand binding protein
ORF01911	SAG1714	95	ribosomal protein S6
ORF01912	SAG1715	374	A/G-specific adenine glycosylase
ORF01913	SAG1716	197	transcriptional regulator, Cro/CI family
ORF01914	SAG1717	104	thioredoxin
ORF01915	SAG1718	166	PAP2 family protein
ORF01916	SAG1719	779	MutS2 family protein
ORF01917	SAG1720	180	conserved hypothetical protein
ORF01918	SAG1721	103	conserved hypothetical protein
ORF01919	SAG1722	297	ribonuclease HIII
ORF01920	SAG1723	197	signal peptidase I
ORF01921	SAG1724	806	helicase, putative
ORF01922	SAG1725	160	conserved hypothetical protein
ORF01923	SAG1726	364	DNA-damage inducible protein P
ORF01924	SAG1727	770	formate acetyltransferase
ORF01925	SAG1728	124	FMN-binding protein
ORF01926	SAG1729	309	conserved hypothetical protein
ORF01927	SAG1730	251	proteinase, putative, degenerate, FRAMESHIFT
ORF01928	SAG1731	298	membrane protein, putative
ORF01929	SAG1732	282	glycerol uptake facilitator protein, putative
ORF01930	SAG1733	150	universal stress protein family
ORF01931	SAG1734	400	transporter, putative
ORF01932	SAG1735	219	transcriptional regulator, Crp/Fnr family
ORF01933	SAG1736	761	X-pro dipeptidyl-peptidase
ORF01934	SAG1737	119	hypothetical protein
ORF01936	SAG1738	326	polyprenyl synthetase family protein
ORF01937	SAG1739	582	ABC transporter, ATP-binding protein CydC
ORF01938	SAG1740	572	ABC transporter, ATP-binding protein CydD
ORF01939	SAG1741	339	cytochrome d ubiquinol oxidase, subunit II
ORF01940	SAG1742	475	cytochrome d oxidase, subunit I
ORF01941	SAG1743	402	pyridine nucleotide-disulphide oxidoreductase family protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF01942	SAG1744	299	prenyltransferase, UbiA family
ORF01943	SAG1745	148	hypothetical protein
ORF01944	SAG1746	35	hypothetical protein
ORF01945	SAG1747	99	conserved hypothetical protein TIGR00103
ORF01946	SAG1748	396	cyclopropane-fatty-acyl-phospholipid synthase
ORF01947	SAG1749	241	transcriptional regulator, merR family
ORF01948	SAG1750	195	exonuclease
ORF01949	SAG1751	178	conserved hypothetical protein
ORF01950	SAG1752	375	conserved hypothetical protein TIGR00275
ORF01951	SAG1753	260	conserved hypothetical protein
ORF01952	SAG1754	89	ribosomal protein S14
ORF01953	SAG1755	38	hypothetical protein
ORF01954	SAG1756	341	conserved hypothetical protein
ORF01957	SAG1757	336	O-sialoglycoprotein endopeptidase family protein
ORF01958	SAG1758	135	ribosomal-protein-alanine acetyltransferase, putative
ORF01960	SAG1759	230	glycoprotease family protein, putative
ORF01961	SAG1760	76	conserved hypothetical protein
ORF01962	SAG1761	559	metallo-beta-lactamase superfamily protein
ORF01963	SAG1762	169	conserved hypothetical protein
ORF01964	SAG1763	448	glutamine synthetase, type I
ORF01965	SAG1764	123	transcriptional regulator GlnR
ORF01967	SAG1765	179	conserved hypothetical protein
ORF01969	SAG1766	398	phosphoglycerate kinase
ORF01970	SAG1767	289	acid phosphatase
ORF01971	SAG1768	336	glyceraldehyde 3-phosphate dehydrogenase
ORF01972	SAG1769	692	translation elongation factor G
ORF01973	SAG1770	156	ribosomal protein S7
ORF01974	SAG1771	137	ribosomal protein S12
ORF01975	SAG1772	270	pur operon repressor
ORF01976	SAG1773	313	HD domain protein
ORF01977	SAG1774	424	conserved hypothetical protein
ORF01978	SAG1775	210	conserved hypothetical protein
ORF01979	SAG1776	220	ribulose-phosphate 3-epimerase
ORF01980	SAG1777	290	conserved hypothetical protein TIGR00157
ORF01981	SAG1778	283	rRNA (guanine-N1-)-methyltransferase, putative
ORF01983	SAG1779	290	dimethyladenosine transferase
ORF01984	SAG1780	163	hypothetical protein
ORF01985	SAG1781	186	primase-related protein
ORF01987	SAG1782	260	deoxyribonuclease, TatD family
ORF01988	SAG1783	90	hypothetical protein
ORF01989	SAG1784	130	hypothetical protein
ORF01990	SAG1785	430	hypothetical protein
ORF01991	SAG1786	130	hypothetical protein
ORF01992	SAG1787	420	dltD protein
ORF01993	SAG1788	79	D-alanyl carrier protein
ORF01994	SAG1789	421	dltB protein
ORF01996	SAG1790	511	D-alanine-activating enzyme
ORF01997	SAG1791	395	sensor histidine kinase
ORF01998	SAG1792	224	DNA-binding response regulator
ORF01999	SAG1793	44	ribosomal protein L34
ORF02000	SAG1794	451	membrane protein, putative
ORF02001	SAG1795	388	transposase, IS30 family, putative
ORF02002	SAG1796	575	amino acid ABC transporter, permease protein
ORF02004	SAG1797	407	amino acid ABC transporter, ATP-binding protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF02005	SAG1798	39	hypothetical protein
ORF02006	SAG1799	792	xylulose-5-phosphate/fructose-6-phosphate phosphoketolase
ORF02007	SAG1800	363	conserved hypothetical protein
ORF02008	SAG1801	559	transcriptional antiterminator, BglG family
ORF02009	SAG1802	253	conserved hypothetical protein
ORF02010	SAG1803	505	carbohydrate kinase, FGGY family
ORF02011	SAG1804	329	hypothetical protein
ORF02012	SAG1805	483	PTS system component, putative
ORF02015	SAG1806	318	glyoxylate reductase, NADH-dependent
ORF02016	SAG1807	339	hypothetical protein
ORF02017	SAG1808	327	sugar binding transcriptional regulator, LacI family
ORF02018	SAG1809	215	transaldolase family protein
ORF02019	SAG1810	238	carbohydrate isomerase, AraD/FucA family
ORF02020	SAG1811	287	hexulose-6-phosphate isomerase, putative
ORF02021	SAG1812	221	hexulose-6-phosphate synthase, putative
ORF02022	SAG1813	161	PTS system, IIA component
ORF02023	SAG1814	92	PTS system, IIB component
ORF02024	SAG1815	479	transport protein SgaT, putative
ORF02025	SAG1816	205	hypothetical protein
ORF02026	SAG1817	157	hypothetical protein
ORF02027	SAG1818	430	adenylosuccinate synthetase
ORF02028	SAG1819	340	perfringolysin O regulator protein
ORF02029	SAG1820	224	conserved hypothetical protein
ORF02030	SAG1821	750	glutamate--cysteine ligase-related protein
ORF02031	SAG1822	272	conserved hypothetical protein
ORF02032	SAG1823	418	conserved hypothetical protein
ORF02033	SAG1824	291	chaperonin, 33 kDa
ORF02034	SAG1825	325	NifR3/Smm1 family protein
ORF02035	SAG1826	213	deoxynucleoside kinase family protein
ORF02036	SAG1827	163	phosphinothricin N-acetyltransferase
ORF02037	SAG1828	815	ATP-dependent Clp protease, ATP-binding subunit
ORF02038	SAG1829	154	transcriptional regulator CtsR
ORF02039	SAG1830	153	conserved hypothetical protein
ORF02040	SAG1831	346	translation elongation factor Ts
ORF02041	SAG1832	256	ribosomal protein S2
ORF02042	SAG1833	186	alkyl hydroperoxide reductase, subunit C
ORF02043	SAG1834	510	alkyl hydroperoxide reductase, subunit F
ORF02044	SAG1835	134	conserved hypothetical protein
ORF02045	SAG1836	61	conserved hypothetical protein
ORF02046	SAG1837	468	lysin, putative
ORF02047	SAG1838	109	holin, putative
ORF02048	SAG1839	136	conserved hypothetical protein
ORF02049	SAG1840	112	hypothetical protein
ORF02050	SAG1841	76	conserved domain protein
ORF02051	SAG1842	1224	PblB, putative
ORF02053	SAG1843	240	conserved hypothetical protein
ORF02056	SAG1844	911	conserved hypothetical protein
ORF02057	SAG1845	42	hypothetical protein
ORF02058	SAG1846	158	hypothetical protein
ORF02059	SAG1847	227	conserved hypothetical protein
ORF02060	SAG1848	114	conserved hypothetical protein
ORF02061	SAG1849	115	hypothetical protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF02062	SAG1850	101	hypothetical protein
ORF02063	SAG1851	111	conserved domain protein
ORF02064	SAG1852	420	conserved domain protein
ORF02066	SAG1853	180	protease, putative
ORF02067	SAG1854	380	conserved hypothetical protein
ORF02068	SAG1855	570	terminase large subunit, putative
ORF02069	SAG1856	161	hypothetical protein
ORF02070	SAG1858	95	hypothetical protein
ORF02071	SAG1859	180	site-specific recombinase, phage integrase family
ORF02072	SAG1860	154	conserved hypothetical protein
ORF02073	SAG1861	119	transcriptional regulator, Cro/CI family
ORF02075	SAG1862	86	hypothetical protein
ORF02076	SAG1863	138	single-strand binding protein
ORF02077	SAG1864	68	hypothetical protein
ORF02078	SAG1865	74	conserved hypothetical protein
ORF02079	SAG1866	109	conserved hypothetical protein
ORF02080	SAG1867	163	conserved hypothetical protein
ORF02081	SAG1868	134	hypothetical protein
ORF02082	SAG1869	437	type II DNA modification methyltransferase, putative
ORF02083	SAG1870	273	DNA replication protein DnaC, putative
ORF02084	SAG1871	248	conserved hypothetical protein
ORF02085	SAG1872	200	hypothetical protein
ORF02086	SAG1873	443	replicative DNA helicase
ORF02087	SAG1874	87	hypothetical protein
ORF02088	SAG1875	94	conserved hypothetical protein
ORF02089	SAG1876	176	HNH endonuclease family protein
ORF02090	SAG1877	236	antirepressor protein, putative
ORF02091	SAG1878	102	conserved domain protein
ORF02092	SAG1879	156	hypothetical protein
ORF02093	SAG1880	54	hypothetical protein
ORF02094	SAG1881	51	hypothetical protein
ORF02095	SAG1882	120	repressor protein, putative
ORF02097	SAG1884	134	hypothetical protein
ORF02098	SAG1885	356	site-specific recombinase, phage integrase family
ORF02100	SAG1886	32	hypothetical protein
ORF02101	SAG1887	689	Na+/H+ exchanger family protein
ORF02102	SAG1888	78	hypothetical protein
ORF02103	SAG1889	317	microcin immunity protein MccF, putative
ORF02104	SAG1890	631	endopeptidase O
ORF02105	SAG1891	327	oxidoreductase, Gfo/Idh/MocA family
ORF02107	SAG1892	358	membrane protein, putative
ORF02108	SAG1893	59	hypothetical protein
ORF02109	SAG1894	214	Cyclic nucleotide-binding domain protein
ORF02110	SAG1895	204	polypeptide deformylase
ORF02111	SAG1896	333	sugar binding transcriptional regulator RegR
ORF02112	SAG1897	634	conserved hypothetical protein
ORF02113	SAG1898	271	PTS system, IID component
ORF02114	SAG1899	288	PTS system, IIC component
ORF02115	SAG1900	164	PTS system, IIB component
ORF02116	SAG1901	398	glucuronyl hydrolase
ORF02118	SAG1902	144	PTS system, IIA component
ORF02119	SAG1903	34	hypothetical protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF02120	SAG1904	270	oxidoreductase, short-chain dehydrogenase/reductase family
ORF02121	SAG1905	212	conserved hypothetical protein
ORF02122	SAG1906	335	carbohydrate kinase, PfkB family
ORF02123	SAG1907	212	2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase
ORF02124	SAG1908	499	hypothetical protein
ORF02125	SAG1909	204	nitroreductase family protein
ORF02126	SAG1910	141	transcriptional regulator, MarR family
ORF02127	SAG1911	1468	DNA polymerase III, alpha subunit, Gram-positive type
ORF02128	SAG1912	194	N-acetylmuramoyl-L-alanine amidase, family 4 protein
ORF02129	SAG1913	617	prolyl-tRNA synthetase
ORF02130	SAG1914	419	membrane-associated zinc metalloprotease, putative
ORF02131	SAG1915	264	phosphatidate cytidylyltransferase
ORF02132	SAG1916	250	undecaprenyl diphosphate synthase
ORF02133	SAG1917	113	preprotein translocase, YajC subunit
ORF02134	SAG1918	114	conserved hypothetical protein
ORF02135	SAG1919	387	malate oxidoreductase
ORF02136	SAG1920	445	citrate carrier protein, CCS family
ORF02137	SAG1921	508	sensor histidine kinase family protein
ORF02138	SAG1922	229	response regulator
ORF02139	SAG1923	331	UDP-glucose 4-epimerase
ORF02140	SAG1924	535	glucan 1,6-alpha-glucosidase
ORF02141	SAG1925	377	sugar ABC transporter, ATP-binding protein
ORF02142	SAG1926	283	helix-turn-helix domain protein, fis-type
ORF02143	SAG1927	298	lacX protein
ORF02144	SAG1928	325	tagatose 1,6-diphosphate aldolase
ORF02145	SAG1929	310	tagatose-6-phosphate kinase
ORF02146	SAG1930	171	galactose-6-phosphate isomerase, LacB subunit
ORF02147	SAG1931	141	galactose-6-phosphate isomerase, LacA subunit
ORF02148	SAG1932	816	neuraminidase
ORF02149	SAG1933	482	PTS system, IIC component, putative
ORF02150	SAG1934	101	PTS system, IIB component, putative
ORF02152	SAG1935	157	PTS system, IIA component, putative
ORF02153	SAG1936	258	lactose phosphotransferase system repressor
ORF02156	SAG1937		streptococcal histidine triad family protein, degenerate, FRAMESHIFT
ORF02157	SAG1938	307	adhesion lipoprotein, putative
ORF02158	SAG1939	147	conserved hypothetical protein TIGR00256
ORF02159	SAG1940	738	GTP pyrophosphokinase
ORF02160	SAG1941	800	2`,3`-cyclic-nucleotide 2`-phosphodiesterase
ORF02161	SAG1942	151	nrdI protein, putative
ORF02162	SAG1943	345	conserved hypothetical protein
ORF02163	SAG1944	165	conserved hypothetical protein
ORF02164	SAG1945	345	iron ABC transporter, iron-binding protein
ORF02165	SAG1946	257	DNA-binding response regulator
ORF02166	SAG1947	549	conserved hypothetical protein
ORF02167	SAG1948	275	PTS system, IID component
ORF02168	SAG1949	269	PTS system, IIC component
ORF02169	SAG1950	163	PTS system, IIB component
ORF02170	SAG1951	141	PTS system, IIA component, putative
ORF02171	SAG1952	353	membrane protein, putative
ORF02172	SAG1953	60	hypothetical protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF02173	SAG1954	384	hypothetical protein
ORF02174	SAG1955	282	ABC transporter, ATP-binding protein
ORF02175	SAG1956	96	conserved domain protein
ORF02176	SAG1957	250	response regulator
ORF02177	SAG1958	276	conserved hypothetical protein
ORF02178	SAG1959	727	PTS system, IIABC components
ORF02179	SAG1960	551	sensor histidine kinase
ORF02180	SAG1961	225	phosphate regulon response regulator PhoB
ORF02181	SAG1962	218	phosphate transport system regulatory protein PhoU, putative
ORF02182	SAG1963	253	phosphate ABC transporter, ATP-binding protein
ORF02183	SAG1964	292	phosphate ABC transporter, permease protein
ORF02184	SAG1965	281	phosphate ABC transporter, permease protein
ORF02186	SAG1966	293	hemolysin precursor, putative
ORF02187	SAG1967	195	hypothetical protein
ORF02188	SAG1968	246	conserved hypothetical protein TIGR00046
ORF02189	SAG1969	317	ribosomal protein L11 methyltransferase
ORF02190	SAG1970	102	conserved hypothetical protein
ORF02191	SAG1971	41	hypothetical protein
ORF02192	SAG1972	238	transcriptional regulator, MerR family
ORF02194	SAG1973	156	acetyltransferase, GNAT family
ORF02195	SAG1974	152	MutT/nudix family protein
ORF02196	SAG1975	47	hypothetical protein
ORF02197	SAG1976	156	conserved hypothetical protein
ORF02198	SAG1977	163	acetyltransferase, GNAT family
ORF02199	SAG1978	422	ATPase, AAA family
ORF02201	SAG1979	253	hypothetical protein
ORF02202	SAG1980	300	ABC transporter, ATP-binding protein
ORF02203	SAG1981	68	hypothetical protein
ORF02205	SAG1982	359	transcriptional regulator, Cro/CI family
ORF02206	SAG1983	105	conserved hypothetical protein
ORF02207	SAG1984	188	conserved hypothetical protein TIGR00730
ORF02208	SAG1985	51	hypothetical protein
ORF02209	SAG1986	375	integrase, phage family, putative
ORF02210	SAG1987	61	conserved hypothetical protein
ORF02211	SAG1988	342	conserved hypothetical protein
ORF02212	SAG1989	139	hypothetical protein
ORF02213	SAG1990	127	hypothetical protein
ORF02214	SAG1991	204	transcriptional regulator, Cro/CI family
ORF02215	SAG1992	518	conserved hypothetical protein
ORF02216	SAG1993	373	site-specific recombinase, phage integrase family
ORF02217	SAG1994	108	conserved hypothetical protein
ORF02219	SAG1995	210	hypothetical protein
ORF02221	SAG1996	263	cell wall anchor protein-related protein
ORF02223	SAG1997	182	hypothetical protein
ORF02224	SAG1998	457	hypothetical protein
ORF02225	SAG1999	47	hypothetical protein
ORF02226	SAG2000	666	membrane protein, putative
ORF02227	SAG2001	756	conjugal transfer protein, interruption-C
ORF02228	SAG2002	129	IS1381, transposase OrfB
ORF02229	SAG2003	127	IS1381, transposase OrfA
ORF02230	SAG2005	136	conserved hypothetical protein
ORF02231	SAG2006	88	conserved hypothetical protein
ORF02232	SAG2007	317	conserved hypothetical protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF02233	SAG2008	84	conserved hypothetical protein
ORF02234	SAG2009	88	conserved hypothetical protein
ORF02235	SAG2010	157	hypothetical protein
ORF02236	SAG2011	160	conserved hypothetical protein
ORF02237	SAG2012	90	hypothetical protein
ORF02238	SAG2013	189	hypothetical protein
ORF02239	SAG2014	449	hypothetical protein
ORF02240	SAG2015	99	transcriptional regulator, Cro/CI family
ORF02241	SAG2016	125	hypothetical protein
ORF02242	SAG2017	429	transcriptional regulator, Cro/CI family
ORF02243	SAG2018	553	FtsK/SpoIIIE family protein
ORF02244	SAG2019	153	hypothetical protein
ORF02245	SAG2020	98	hypothetical protein
ORF02246	SAG2021	826	cell wall surface anchor family protein
ORF02247	SAG2022	417	transposase, ISL3 family
ORF02249	SAG2023	546	mercuric reductase
ORF02250	SAG2024	130	mercuric resistance operon regulatory protein MerR
ORF02251	SAG2025	522	Mn2+/Fe2+ transporter, NRAMP family
ORF02252	SAG2026	240	membrane protein, putative
ORF02253	SAG2027	205	ABC transporter, ATP-binding protein
ORF02254	SAG2028	36	conserved hypothetical protein
ORF02255	SAG2029	284	streptomycin resistance protein
ORF02257	SAG2030	130	hypothetical protein
ORF02258	SAG2031	202	hypothetical protein
ORF02259	SAG2032	111	conserved hypothetical protein
ORF02260	SAG2033	162	acetyltransferase, GNAT family
ORF02261	SAG2034	247	membrane protein, putative
ORF02262	SAG2035	300	ABC transporter, ATP-binding protein
ORF02263	SAG2036	68	hypothetical protein
ORF02264	SAG2037	358	transcriptional regulator, Cro/CI family
ORF02265	SAG2038	204	PAP2 family protein
ORF02266	SAG2039	98	conserved hypothetical protein
ORF02267	SAG2040	186	conserved hypothetical protein TIGR00730
ORF02268	SAG2041	287	protease, putative
ORF02269	SAG2042	100	rhodanese family protein
ORF02270	SAG2043	255	cAMP factor
ORF02271	SAG2044	62	hypothetical protein
ORF02272	SAG2045	179	DNA topology modulation protein FlaR, putative
ORF02273	SAG2046	361	glycerol dehydrogenase, putative
ORF02274	SAG2047	235	conserved hypothetical protein
ORF02275	SAG2048	614	5-methyltetrahydrofolate--homocysteine methyltransferase, putative
ORF02276	SAG2049	745	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase
ORF02277	SAG2050	107	conserved hypothetical protein
ORF02278	SAG2051	230	branched-chain amino acid transport protein AzlC, putative
ORF02279	SAG2052	41	hypothetical protein
ORF02280	SAG2053	1570	serine protease, subtilase family, putative
ORF02281	SAG2054	228	DNA-binding response regulator
ORF02282	SAG2055	462	sensor histidine kinase
ORF02283	SAG2056	202	chromosome assembly-related protein
ORF02285	SAG2057	833	leucyl-tRNA synthetase
ORF02286	SAG2058	415	major facilitator family protein

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF02287	SAG2059	281	conserved hypothetical protein
ORF02288	SAG2060	398	glycosyl transferase, family 8
ORF02289	SAG2061	401	glycosyl transferase, family 8
ORF02290	SAG2062	179	transcription antitermination protein NusG
ORF02291	SAG2063	630	pathogenicity protein, putative
ORF02292	SAG2064	57	preprotein translocase, SecE subunit, putative
ORF02293	SAG2066	773	penicillin-binding protein 2A
ORF02294	SAG2067	294	ribosomal large subunit pseudouridine synthase, RluD subfamily
ORF02295	SAG2068	546	Lyme disease proteins of unknown function, putative
ORF02296	SAG2069	403	phosphopentomutase
ORF02297	SAG2070	223	deoxyribose-phosphate aldolase
ORF02298	SAG2071	400	Na+ dependent nucleoside transporter
ORF02300	SAG2072	259	uridine phosphorylase
ORF02301	SAG2073	245	transcriptional regulator, GntR family
ORF02302	SAG2074	540	60 kda chaperonin
ORF02303	SAG2075	94	chaperonin, 10 kDa
ORF02305	SAG2076	267	ABC transporter, ATP-binding protein
ORF02306	SAG2077	298	ABC transporter, permease protein
ORF02307	SAG2078	320	lipoprotein, putative
ORF02308	SAG2079	265	hydrolase, haloacid dehalogenase-like family
ORF02309	SAG2080	286	glyoxalase family protein
ORF02310	SAG2081	243	conserved hypothetical protein
ORF02311	SAG2082	205	anaerobic ribonucleoside-triphosphate reductase activating protein
ORF02312	SAG2083	163	acetyltransferase, GNAT family
ORF02313	SAG2084	310	virulence factor MviM, putative
ORF02314	SAG2085	47	conserved hypothetical protein
ORF02315	SAG2086	723	anaerobic ribonucleoside-triphosphate reductase
ORF02316	SAG2087	495	conserved hypothetical protein
ORF02317	SAG2088	40	hypothetical protein
ORF02318	SAG2089	105	conserved hypothetical protein
ORF02319	SAG2090	136	conserved hypothetical protein TIGR00250
ORF02320	SAG2091	88	conserved hypothetical protein
ORF02321	SAG2092	132	conserved hypothetical protein
ORF02322	SAG2093	379	recA protein
ORF02323	SAG2094		competence/damage-inducible protein CinA FRAMESHIFT
ORF02325	SAG2095	183	DNA-3-methyladenine glycosylase I
ORF02327	SAG2096	196	Holliday junction DNA helicase RuvA
ORF02328	SAG2097	418	transporter, putative
ORF02329	SAG2098	659	DNA mismatch repair protein HexB
ORF02330	SAG2099	33	hypothetical protein
ORF02331	SAG2100	67	cold shock protein, CSD family
ORF02332	SAG2101	858	DNA mismatch repair protein HexA
ORF02333	SAG2102	145	arginine repressor ArgR, putative
ORF02334	SAG2103	563	arginyl-tRNA synthetase
ORF02335	SAG2104	102	conserved hypothetical protein
ORF02337	SAG2105	290	conserved hypothetical protein
ORF02338	SAG2106	314	conserved hypothetical protein
ORF02339	SAG2107	583	aspartyl-tRNA synthetase
ORF02340	SAG2108	426	histidyl-tRNA synthetase
ORF02341	SAG2109	60	ribosomal protein L32

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF02342	SAG2110	49	ribosomal protein L33
ORF02343	SAG2111	173	conserved hypothetical protein
ORF02344	SAG2112	494	site-specific recombinase, phage integrase family
ORF02345	SAG2113	82	conserved hypothetical protein
ORF02346	SAG2114	342	conserved hypothetical protein
ORF02347	SAG2115	143	hypothetical protein
ORF02349	SAG2116	151	conserved hypothetical protein
ORF02350	SAG2117	71	hypothetical protein
ORF02351	SAG2118	306	transcriptional regulator, Cro/CI family
ORF02352	SAG2119	373	conserved domain protein
ORF02355	SAG2120	56	hypothetical protein
ORF02356	SAG2121	176	hypothetical protein
ORF02357	SAG2122	223	DNA-binding response regulator
ORF02358	SAG2123	454	sensor histidine kinase
ORF02359	SAG2124	517	membrane protein, putative
ORF02360	SAG2125	308	carbamate kinase
ORF02361	SAG2126	332	ornithine carbamoyltransferase
ORF02362	SAG2127	431	sensor histidine kinase
ORF02363	SAG2128	277	response regulator
ORF02364	SAG2129	240	amino acid ABC transporter, ATP-binding protein
ORF02365	SAG2130	504	amino acid ABC transporter, permease and amino acid binding protein
ORF02367	SAG2131	847	membrane protein, putative
ORF02368	SAG2132	247	conserved hypothetical protein
ORF02369	SAG2133	118	conserved hypothetical protein
ORF02370	SAG2134	772	membrane protein, putative
ORF02371	SAG2135	179	transcriptional regulator, TetR family, putative
ORF02372	SAG2136	98	conserved hypothetical protein
ORF02373	SAG2137	203	ribosomal protein S4
ORF02374	SAG2138	95	conserved hypothetical protein
ORF02375	SAG2139	451	replicative DNA helicase
ORF02376	SAG2140	150	ribosomal protein L9
ORF02377	SAG2141	660	DHH family protein
ORF02378	SAG2142	613	glucose inhibited division protein A
ORF02379	SAG2143	203	conserved hypothetical protein TIGR00427
ORF02380	SAG2144	373	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
ORF02381	SAG2145	222	L-serine dehydratase, iron-sulfur-dependent, beta subunit
ORF02382	SAG2146	290	L-serine dehydratase, iron-sulfur-dependent, alpha subunit
ORF02383	SAG2147	234	conserved hypothetical protein
ORF02384	SAG2148	179	LysM domain protein
ORF02385	SAG2149	264	cobalt transport family protein
ORF02386	SAG2150	280	ABC transporter, ATP-binding protein
ORF02387	SAG2151	279	ABC transporter, ATP-binding protein, FRAMESHIFT
ORF02388	SAG2152	180	CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase
ORF02389	SAG2153	427	peptidase, M16 family
ORF02390	SAG2154	414	conserved hypothetical protein
ORF02391	SAG2155	117	conserved hypothetical protein
ORF02392	SAG2156	369	recF protein
ORF02393	SAG2157	278	transporter, putative
ORF02395	SAG2158	220	transcriptional regulator, Cro/CI family

Table 32: Conversion of ORF Ref Nos. with SAG Ref Nos.

ORF Ref No.	SAGxxxx Ref No.	aa	Annotation
ORF02396	SAG2159	493	inosine-5'-monophosphate dehydrogenase
ORF02397	SAG2160	161	transcriptional regulator, ArgR family
ORF02398	SAG2161	226	transcriptional regulator, Crp/Fnr family
ORF02399	SAG2162	234	conserved hypothetical protein
ORF02400	SAG2163	410	arginine deiminase
ORF02401	SAG2164	136	acetyltransferase, GNAT family
ORF02402	SAG2165	337	ornithine carbamoyltransferase
ORF02403	SAG2166	475	arginine/ornithine antiporter
ORF02404	SAG2167	318	carbamate kinase
ORF02405	SAG2168	341	tryptophanyl-tRNA synthetase
ORF02406	SAG2169	230	conserved hypothetical protein
ORF02407	SAG2170	290	conserved hypothetical protein
ORF02408	SAG2171	539	ABC transporter, ATP-binding protein
ORF02409	SAG2172	859	ABC transporter, permease protein, putative
ORF02410	SAG2173	159	conserved hypothetical protein TIGR00246
ORF02411	SAG2174	409	serine protease
ORF02412	SAG2175	257	partitioning protein, ParB family
ORF02413	SAG0001	453	chromosomal replication initiator protein DnaA
ORF02415	SAG0002	378	DNA polymerase III, beta subunit
ORF02416	SAG0003	293	diacylglycerol kinase catalytic domain protein, putative
ORF02417	SAG0004	65	conserved hypothetical protein
ORF02418	SAG0005	67	hypothetical protein
ORF02419	SAG0006	371	conserved hypothetical GTP-binding protein
ORF02420	SAG0007	191	peptidyl-tRNA hydrolase
ORF02421	SAG0008	1165	transcription-repair coupling factor
ORF02422	SAG0009	31	hypothetical protein
ORF02423	SAG0010	90	S4 domain protein
ORF02424	SAG0011	123	cell division protein DivIC, putative
ORF02425	SAG0012	44	conserved hypothetical protein
ORF02426	SAG0013	428	conserved hypothetical protein
ORF02427	SAG0014	424	MesJ/Ycf62 family protein
ORF02428	SAG0015	180	hypoxanthine-guanine phosphoribosyltransferase
ORF02429	SAG0016	658	cell division protein FtsH
ORF03000	SAG0157		Dnase-related protein, DEGENERATE
ORF03001	SAG0579	142	conserved hypothetical protein
ORF03002	SAG0580	111	conserved hypothetical protein, truncation
ORF03003	SAG0652		Tn5252, Orf 28 protein, degenerate
ORF03004	SAG0655	57	conserved hypothetical protein
ORF03005	SAG0662	101	cylX protein
ORF03006	SAG0917	83	Tn916, hypothetical protein
ORF03007	SAG0920	23	Tn916, hypothetical protein
ORF03008	SAG0922	61	Tn916, hypothetical protein
ORF03009	SAG0924	28	Tn916, tetM leader peptide
ORF03010	SAG0936	39	Tn916, hypothetical protein
ORF03011	SAG1484	48	ribosomal protein L33
ORF03012	SAG1857	119	HNH endonuclease family protein
ORF03013	SAG1883	128	conserved hypothetical protein
ORF03014	SAG2065	50	ribosomal protein L33
ORF03015	SAG2004	67	conjugal transfer protein, interruption-N

Table 33: List of GAS ORFs which are shared with GBS and Spn

gi|13621326|gb|AAK33146.1|
gi|13621327|gb|AAK33147.1|
gi|13621328|gb|AAK33148.1|
gi|13621329|gb|AAK33149.1|
gi|13621330|gb|AAK33150.1|
gi|13621331|gb|AAK33151.1|
gi|13621332|gb|AAK33152.1|
gi|13621333|gb|AAK33153.1|
gi|13621334|gb|AAK33154.1|
gi|13621335|gb|AAK33155.1|
gi|13621337|gb|AAK33156.1|
gi|13621340|gb|AAK33158.1|
gi|13621341|gb|AAK33159.1|
gi|13621343|gb|AAK33160.1|
gi|13621344|gb|AAK33161.1|
gi|13621346|gb|AAK33163.1|
gi|13621347|gb|AAK33164.1|
gi|13621348|gb|AAK33165.1|
gi|13621349|gb|AAK33166.1|
gi|13621350|gb|AAK33167.1|
gi|13621353|gb|AAK33169.1|
gi|13621354|gb|AAK33170.1|
gi|13621355|gb|AAK33171.1|
gi|13621357|gb|AAK33173.1|
gi|13621358|gb|AAK33174.1|
gi|13621359|gb|AAK33175.1|
gi|13621361|gb|AAK33176.1|
gi|13621362|gb|AAK33177.1|
gi|13621363|gb|AAK33178.1|
gi|13621364|gb|AAK33179.1|
gi|13621365|gb|AAK33180.1|
gi|13621366|gb|AAK33181.1|
gi|13621367|gb|AAK33182.1|
gi|13621368|gb|AAK33183.1|
gi|13621369|gb|AAK33184.1|
gi|13621370|gb|AAK33185.1|
gi|13621372|gb|AAK33186.1|
gi|13621373|gb|AAK33187.1|
gi|13621374|gb|AAK33188.1|
gi|13621375|gb|AAK33189.1|
gi|13621376|gb|AAK33190.1|
gi|13621377|gb|AAK33191.1|
gi|13621378|gb|AAK33192.1|
gi|13621379|gb|AAK33193.1|
gi|13621380|gb|AAK33194.1|
gi|13621382|gb|AAK33196.1|
gi|13621383|gb|AAK33197.1|
gi|13621384|gb|AAK33198.1|
gi|13621385|gb|AAK33199.1|
gi|13621386|gb|AAK33200.1|
gi|13621387|gb|AAK33201.1|
gi|13621388|gb|AAK33202.1|
gi|13621389|gb|AAK33203.1|
gi|13621390|gb|AAK33204.1|
gi|13621391|gb|AAK33205.1|
gi|13621392|gb|AAK33206.1|
gi|13621393|gb|AAK33207.1|
gi|13621394|gb|AAK33208.1|
gi|13621397|gb|AAK33210.1|
gi|13621398|gb|AAK33211.1|
gi|13621399|gb|AAK33212.1|
gi|13621401|gb|AAK33214.1|
gi|13621403|gb|AAK33215.1|
gi|13621404|gb|AAK33216.1|
gi|13621405|gb|AAK33217.1|
gi|13621407|gb|AAK33218.1|
gi|13621408|gb|AAK33219.1|
gi|13621409|gb|AAK33220.1|
gi|13621413|gb|AAK33224.1|
gi|13621415|gb|AAK33226.1|
gi|13621416|gb|AAK33227.1|
gi|13621418|gb|AAK33229.1|
gi|13621419|gb|AAK33230.1|
gi|13621424|gb|AAK33234.1|
gi|13621425|gb|AAK33235.1|
gi|13621426|gb|AAK33236.1|
gi|13621434|gb|AAK33243.1|
gi|13621450|gb|AAK33258.1|
gi|13621455|gb|AAK33262.1|
gi|13621456|gb|AAK33263.1|
gi|13621457|gb|AAK33264.1|
gi|13621467|gb|AAK33273.1|
gi|13621468|gb|AAK33274.1|
gi|13621469|gb|AAK33275.1|
gi|13621470|gb|AAK33276.1|
gi|13621471|gb|AAK33277.1|
gi|13621472|gb|AAK33278.1|
gi|13621473|gb|AAK33279.1|
gi|13621476|gb|AAK33281.1|
gi|13621477|gb|AAK33282.1|
gi|13621478|gb|AAK33283.1|
gi|13621480|gb|AAK33285.1|
gi|13621481|gb|AAK33286.1|
gi|13621491|gb|AAK33295.1|
gi|13621494|gb|AAK33298.1|
gi|13621496|gb|AAK33299.1|
gi|13621501|gb|AAK33304.1|
gi|13621502|gb|AAK33305.1|
gi|13621505|gb|AAK33307.1|
gi|13621506|gb|AAK33308.1|
gi|13621507|gb|AAK33309.1|
gi|13621510|gb|AAK33312.1|
gi|13621511|gb|AAK33313.1|
gi|13621513|gb|AAK33315.1|
gi|13621516|gb|AAK33317.1|
gi|13621518|gb|AAK33319.1|
gi|13621521|gb|AAK33322.1|
gi|13621522|gb|AAK33323.1|
gi|13621523|gb|AAK33324.1|
gi|13621524|gb|AAK33325.1|
gi|13621525|gb|AAK33326.1|
gi|13621527|gb|AAK33327.1|

Table 33: List of GAS ORFs which are shared with GBS and Spn

gi|13621528|gb|AAK33328.1|
gi|13621529|gb|AAK33329.1|
gi|13621530|gb|AAK33330.1|
gi|13621531|gb|AAK33331.1|
gi|13621532|gb|AAK33332.1|
gi|13621533|gb|AAK33333.1|
gi|13621534|gb|AAK33334.1|
gi|13621535|gb|AAK33335.1|
gi|13621536|gb|AAK33336.1|
gi|13621537|gb|AAK33337.1|
gi|13621539|gb|AAK33338.1|
gi|13621540|gb|AAK33339.1|
gi|13621541|gb|AAK33340.1|
gi|13621542|gb|AAK33341.1|
gi|13621543|gb|AAK33342.1|
gi|13621544|gb|AAK33343.1|
gi|13621546|gb|AAK33345.1|
gi|13621547|gb|AAK33346.1|
gi|13621548|gb|AAK33347.1|
gi|13621550|gb|AAK33348.1|
gi|13621551|gb|AAK33349.1|
gi|13621552|gb|AAK33350.1|
gi|13621553|gb|AAK33351.1|
gi|13621554|gb|AAK33352.1|
gi|13621555|gb|AAK33353.1|
gi|13621557|gb|AAK33355.1|
gi|13621559|gb|AAK33356.1|
gi|13621560|gb|AAK33357.1|
gi|13621561|gb|AAK33358.1|
gi|13621562|gb|AAK33359.1|
gi|13621563|gb|AAK33360.1|
gi|13621564|gb|AAK33361.1|
gi|13621565|gb|AAK33362.1|
gi|13621566|gb|AAK33363.1|
gi|13621567|gb|AAK33364.1|
gi|13621569|gb|AAK33365.1|
gi|13621571|gb|AAK33367.1|
gi|13621572|gb|AAK33368.1|
gi|13621573|gb|AAK33369.1|
gi|13621574|gb|AAK33370.1|
gi|13621575|gb|AAK33371.1|
gi|13621576|gb|AAK33372.1|
gi|13621577|gb|AAK33373.1|
gi|13621579|gb|AAK33374.1|
gi|13621581|gb|AAK33376.1|
gi|13621582|gb|AAK33377.1|
gi|13621583|gb|AAK33378.1|
gi|13621584|gb|AAK33379.1|
gi|13621585|gb|AAK33380.1|
gi|13621586|gb|AAK33381.1|
gi|13621588|gb|AAK33383.1|
gi|13621589|gb|AAK33384.1|
gi|13621590|gb|AAK33385.1|
gi|13621592|gb|AAK33386.1|
gi|13621593|gb|AAK33387.1|
gi|13621594|gb|AAK33388.1|
gi|13621595|gb|AAK33389.1|
gi|13621596|gb|AAK33390.1|
gi|13621597|gb|AAK33391.1|
gi|13621598|gb|AAK33392.1|
gi|13621599|gb|AAK33393.1|
gi|13621600|gb|AAK33394.1|
gi|13621602|gb|AAK33395.1|
gi|13621603|gb|AAK33396.1|
gi|13621604|gb|AAK33397.1|
gi|13621605|gb|AAK33398.1|
gi|13621606|gb|AAK33399.1|
gi|13621607|gb|AAK33400.1|
gi|13621608|gb|AAK33401.1|
gi|13621609|gb|AAK33402.1|
gi|13621611|gb|AAK33404.1|
gi|13621614|gb|AAK33406.1|
gi|13621615|gb|AAK33407.1|
gi|13621616|gb|AAK33408.1|
gi|13621617|gb|AAK33409.1|
gi|13621618|gb|AAK33410.1|
gi|13621619|gb|AAK33411.1|
gi|13621620|gb|AAK33412.1|
gi|13621621|gb|AAK33413.1|
gi|13621622|gb|AAK33414.1|
gi|13621623|gb|AAK33415.1|
gi|13621624|gb|AAK33416.1|
gi|13621625|gb|AAK33417.1|
gi|13621627|gb|AAK33419.1|
gi|13621629|gb|AAK33420.1|
gi|13621630|gb|AAK33421.1|
gi|13621631|gb|AAK33422.1|
gi|13621633|gb|AAK33424.1|
gi|13621634|gb|AAK33425.1|
gi|13621636|gb|AAK33427.1|
gi|13621637|gb|AAK33428.1|
gi|13621638|gb|AAK33429.1|
gi|13621640|gb|AAK33430.1|
gi|13621642|gb|AAK33432.1|
gi|13621644|gb|AAK33434.1|
gi|13621645|gb|AAK33435.1|
gi|13621647|gb|AAK33437.1|
gi|13621648|gb|AAK33438.1|
gi|13621650|gb|AAK33440.1|
gi|13621651|gb|AAK33441.1|
gi|13621652|gb|AAK33442.1|
gi|13621657|gb|AAK33446.1|
gi|13621658|gb|AAK33447.1|
gi|13621660|gb|AAK33449.1|
gi|13621670|gb|AAK33458.1|
gi|13621671|gb|AAK33459.1|
gi|13621672|gb|AAK33460.1|
gi|13621675|gb|AAK33462.1|
gi|13621676|gb|AAK33463.1|
gi|13621678|gb|AAK33465.1|
gi|13621680|gb|AAK33467.1|
gi|13621681|gb|AAK33468.1|

Table 33: List of GAS ORFs which are shared with GBS and Spn

gi|13621682|gb|AAK33469.1|
gi|13621683|gb|AAK33470.1|
gi|13621684|gb|AAK33471.1|
gi|13621685|gb|AAK33472.1|
gi|13621688|gb|AAK33474.1|
gi|13621689|gb|AAK33475.1|
gi|13621690|gb|AAK33476.1|
gi|13621691|gb|AAK33477.1|
gi|13621692|gb|AAK33478.1|
gi|13621693|gb|AAK33479.1|
gi|13621694|gb|AAK33480.1|
gi|13621695|gb|AAK33481.1|
gi|13621697|gb|AAK33483.1|
gi|13621698|gb|AAK33484.1|
gi|13621700|gb|AAK33485.1|
gi|13621701|gb|AAK33486.1|
gi|13621702|gb|AAK33487.1|
gi|13621714|gb|AAK33498.1|
gi|13621715|gb|AAK33499.1|
gi|13621717|gb|AAK33501.1|
gi|13621718|gb|AAK33502.1|
gi|13621719|gb|AAK33503.1|
gi|13621720|gb|AAK33504.1|
gi|13621726|gb|AAK33509.1|
gi|13621727|gb|AAK33510.1|
gi|13621729|gb|AAK33512.1|
gi|13621730|gb|AAK33513.1|
gi|13621731|gb|AAK33514.1|
gi|13621732|gb|AAK33515.1|
gi|13621733|gb|AAK33516.1|
gi|13621734|gb|AAK33517.1|
gi|13621735|gb|AAK33518.1|
gi|13621736|gb|AAK33519.1|
gi|13621741|gb|AAK33523.1|
gi|13621742|gb|AAK33524.1|
gi|13621743|gb|AAK33525.1|
gi|13621744|gb|AAK33526.1|
gi|13621745|gb|AAK33527.1|
gi|13621747|gb|AAK33528.1|
gi|13621756|gb|AAK33537.1|
gi|13621773|gb|AAK33552.1|
gi|13621774|gb|AAK33553.1|
gi|13621775|gb|AAK33554.1|
gi|13621777|gb|AAK33556.1|
gi|13621778|gb|AAK33557.1|
gi|13621779|gb|AAK33558.1|
gi|13621781|gb|AAK33559.1|
gi|13621782|gb|AAK33560.1|
gi|13621785|gb|AAK33563.1|
gi|13621786|gb|AAK33564.1|
gi|13621787|gb|AAK33565.1|
gi|13621788|gb|AAK33566.1|
gi|13621789|gb|AAK33567.1|
gi|13621790|gb|AAK33568.1|
gi|13621793|gb|AAK33571.1|
gi|13621794|gb|AAK33572.1|
gi|13621796|gb|AAK33573.1|
gi|13621797|gb|AAK33574.1|
gi|13621799|gb|AAK33576.1|
gi|13621800|gb|AAK33577.1|
gi|13621802|gb|AAK33579.1|
gi|13621806|gb|AAK33583.1|
gi|13621808|gb|AAK33584.1|
gi|13621809|gb|AAK33585.1|
gi|13621810|gb|AAK33586.1|
gi|13621811|gb|AAK33587.1|
gi|13621812|gb|AAK33588.1|
gi|13621813|gb|AAK33589.1|
gi|13621814|gb|AAK33590.1|
gi|13621817|gb|AAK33592.1|
gi|13621818|gb|AAK33593.1|
gi|13621819|gb|AAK33594.1|
gi|13621820|gb|AAK33595.1|
gi|13621821|gb|AAK33596.1|
gi|13621822|gb|AAK33597.1|
gi|13621823|gb|AAK33598.1|
gi|13621824|gb|AAK33599.1|
gi|13621825|gb|AAK33600.1|
gi|13621826|gb|AAK33601.1|
gi|13621828|gb|AAK33602.1|
gi|13621829|gb|AAK33603.1|
gi|13621830|gb|AAK33604.1|
gi|13621831|gb|AAK33605.1|
gi|13621834|gb|AAK33608.1|
gi|13621835|gb|AAK33609.1|
gi|13621836|gb|AAK33610.1|
gi|13621837|gb|AAK33611.1|
gi|13621839|gb|AAK33612.1|
gi|13621840|gb|AAK33613.1|
gi|13621841|gb|AAK33614.1|
gi|13621842|gb|AAK33615.1|
gi|13621843|gb|AAK33616.1|
gi|13621844|gb|AAK33617.1|
gi|13621898|gb|AAK33667.1|
gi|13621901|gb|AAK33670.1|
gi|13621902|gb|AAK33671.1|
gi|13621904|gb|AAK33672.1|
gi|13621907|gb|AAK33675.1|
gi|13621908|gb|AAK33676.1|
gi|13621909|gb|AAK33677.1|
gi|13621910|gb|AAK33678.1|
gi|13621912|gb|AAK33680.1|
gi|13621924|gb|AAK33690.1|
gi|13621929|gb|AAK33694.1|
gi|13621930|gb|AAK33695.1|
gi|13621931|gb|AAK33696.1|
gi|13621933|gb|AAK33698.1|
gi|13621934|gb|AAK33699.1|
gi|13621935|gb|AAK33700.1|
gi|13621936|gb|AAK33701.1|
gi|13621937|gb|AAK33702.1|
gi|13621938|gb|AAK33703.1|

Table 33: List of GAS ORFs which are shared with GBS and Spn

gi|13621939|gb|AAK33704.1|
gi|13621942|gb|AAK33706.1|
gi|13621944|gb|AAK33708.1|
gi|13621945|gb|AAK33709.1|
gi|13621946|gb|AAK33710.1|
gi|13621950|gb|AAK33714.1|
gi|13621953|gb|AAK33716.1|
gi|13621954|gb|AAK33717.1|
gi|13621955|gb|AAK33718.1|
gi|13621956|gb|AAK33719.1|
gi|13621957|gb|AAK33720.1|
gi|13621958|gb|AAK33721.1|
gi|13621959|gb|AAK33722.1|
gi|13621961|gb|AAK33723.1|
gi|13621975|gb|AAK33736.1|
gi|13621977|gb|AAK33738.1|
gi|13621978|gb|AAK33739.1|
gi|13621979|gb|AAK33740.1|
gi|13621980|gb|AAK33741.1|
gi|13621981|gb|AAK33742.1|
gi|13621982|gb|AAK33743.1|
gi|13621985|gb|AAK33745.1|
gi|13621986|gb|AAK33746.1|
gi|13621987|gb|AAK33747.1|
gi|13621989|gb|AAK33749.1|
gi|13621990|gb|AAK33750.1|
gi|13621992|gb|AAK33752.1|
gi|13621993|gb|AAK33753.1|
gi|13621994|gb|AAK33754.1|
gi|13621996|gb|AAK33755.1|
gi|13621997|gb|AAK33756.1|
gi|13621998|gb|AAK33757.1|
gi|13621999|gb|AAK33758.1|
gi|13622000|gb|AAK33759.1|
gi|13622001|gb|AAK33760.1|
gi|13622002|gb|AAK33761.1|
gi|13622003|gb|AAK33762.1|
gi|13622004|gb|AAK33763.1|
gi|13622005|gb|AAK33764.1|
gi|13622006|gb|AAK33765.1|
gi|13622008|gb|AAK33766.1|
gi|13622009|gb|AAK33767.1|
gi|13622010|gb|AAK33768.1|
gi|13622012|gb|AAK33770.1|
gi|13622013|gb|AAK33771.1|
gi|13622017|gb|AAK33774.1|
gi|13622018|gb|AAK33775.1|
gi|13622019|gb|AAK33776.1|
gi|13622020|gb|AAK33777.1|
gi|13622021|gb|AAK33778.1|
gi|13622024|gb|AAK33781.1|
gi|13622025|gb|AAK33782.1|
gi|13622026|gb|AAK33783.1|
gi|13622031|gb|AAK33787.1|
gi|13622032|gb|AAK33788.1|
gi|13622033|gb|AAK33789.1|
gi|13622034|gb|AAK33790.1|
gi|13622035|gb|AAK33791.1|
gi|13622039|gb|AAK33794.1|
gi|13622041|gb|AAK33796.1|
gi|13622042|gb|AAK33797.1|
gi|13622043|gb|AAK33798.1|
gi|13622044|gb|AAK33799.1|
gi|13622045|gb|AAK33800.1|
gi|13622046|gb|AAK33801.1|
gi|13622048|gb|AAK33802.1|
gi|13622049|gb|AAK33803.1|
gi|13622050|gb|AAK33804.1|
gi|13622051|gb|AAK33805.1|
gi|13622052|gb|AAK33806.1|
gi|13622054|gb|AAK33808.1|
gi|13622055|gb|AAK33809.1|
gi|13622056|gb|AAK33810.1|
gi|13622058|gb|AAK33812.1|
gi|13622060|gb|AAK33813.1|
gi|13622062|gb|AAK33815.1|
gi|13622064|gb|AAK33817.1|
gi|13622065|gb|AAK33818.1|
gi|13622068|gb|AAK33821.1|
gi|13622069|gb|AAK33822.1|
gi|13622070|gb|AAK33823.1|
gi|13622071|gb|AAK33824.1|
gi|13622073|gb|AAK33825.1|
gi|13622074|gb|AAK33826.1|
gi|13622075|gb|AAK33827.1|
gi|13622077|gb|AAK33829.1|
gi|13622079|gb|AAK33831.1|
gi|13622083|gb|AAK33834.1|
gi|13622085|gb|AAK33836.1|
gi|13622086|gb|AAK33837.1|
gi|13622087|gb|AAK33838.1|
gi|13622088|gb|AAK33839.1|
gi|13622089|gb|AAK33840.1|
gi|13622090|gb|AAK33841.1|
gi|13622091|gb|AAK33842.1|
gi|13622092|gb|AAK33843.1|
gi|13622093|gb|AAK33844.1|
gi|13622095|gb|AAK33845.1|
gi|13622096|gb|AAK33846.1|
gi|13622097|gb|AAK33847.1|
gi|13622162|gb|AAK33908.1|
gi|13622163|gb|AAK33909.1|
gi|13622164|gb|AAK33910.1|
gi|13622165|gb|AAK33911.1|
gi|13622166|gb|AAK33912.1|
gi|13622169|gb|AAK33914.1|
gi|13622170|gb|AAK33915.1|
gi|13622171|gb|AAK33916.1|
gi|13622172|gb|AAK33917.1|
gi|13622174|gb|AAK33919.1|
gi|13622175|gb|AAK33920.1|
gi|13622176|gb|AAK33921.1|

Table 33: List of GAS ORFs which are shared with GBS and Spn

gi|13622177|gb|AAK33922.1|
gi|13622179|gb|AAK33923.1|
gi|13622180|gb|AAK33924.1|
gi|13622181|gb|AAK33925.1|
gi|13622182|gb|AAK33926.1|
gi|13622183|gb|AAK33927.1|
gi|13622184|gb|AAK33928.1|
gi|13622185|gb|AAK33929.1|
gi|13622186|gb|AAK33930.1|
gi|13622189|gb|AAK33932.1|
gi|13622190|gb|AAK33933.1|
gi|13622191|gb|AAK33934.1|
gi|13622192|gb|AAK33935.1|
gi|13622198|gb|AAK33940.1|
gi|13622200|gb|AAK33942.1|
gi|13622201|gb|AAK33943.1|
gi|13622204|gb|AAK33946.1|
gi|13622205|gb|AAK33947.1|
gi|13622207|gb|AAK33949.1|
gi|13622208|gb|AAK33950.1|
gi|13622211|gb|AAK33952.1|
gi|13622213|gb|AAK33954.1|
gi|13622214|gb|AAK33955.1|
gi|13622215|gb|AAK33956.1|
gi|13622216|gb|AAK33957.1|
gi|13622217|gb|AAK33958.1|
gi|13622218|gb|AAK33959.1|
gi|13622219|gb|AAK33960.1|
gi|13622222|gb|AAK33962.1|
gi|13622223|gb|AAK33963.1|
gi|13622224|gb|AAK33964.1|
gi|13622233|gb|AAK33972.1|
gi|13622235|gb|AAK33974.1|
gi|13622236|gb|AAK33975.1|
gi|13622237|gb|AAK33976.1|
gi|13622239|gb|AAK33978.1|
gi|13622240|gb|AAK33979.1|
gi|13622241|gb|AAK33980.1|
gi|13622242|gb|AAK33981.1|
gi|13622243|gb|AAK33982.1|
gi|13622244|gb|AAK33983.1|
gi|13622250|gb|AAK33988.1|
gi|13622252|gb|AAK33990.1|
gi|13622253|gb|AAK33991.1|
gi|13622255|gb|AAK33993.1|
gi|13622256|gb|AAK33994.1|
gi|13622257|gb|AAK33995.1|
gi|13622259|gb|AAK33996.1|
gi|13622260|gb|AAK33997.1|
gi|13622261|gb|AAK33998.1|
gi|13622262|gb|AAK33999.1|
gi|13622263|gb|AAK34000.1|
gi|13622264|gb|AAK34001.1|
gi|13622265|gb|AAK34002.1|
gi|13622266|gb|AAK34003.1|
gi|13622268|gb|AAK34005.1|
gi|13622269|gb|AAK34006.1|
gi|13622271|gb|AAK34007.1|
gi|13622272|gb|AAK34008.1|
gi|13622273|gb|AAK34009.1|
gi|13622274|gb|AAK34010.1|
gi|13622275|gb|AAK34011.1|
gi|13622276|gb|AAK34012.1|
gi|13622277|gb|AAK34013.1|
gi|13622278|gb|AAK34014.1|
gi|13622279|gb|AAK34015.1|
gi|13622281|gb|AAK34017.1|
gi|13622282|gb|AAK34018.1|
gi|13622283|gb|AAK34019.1|
gi|13622284|gb|AAK34020.1|
gi|13622285|gb|AAK34021.1|
gi|13622287|gb|AAK34022.1|
gi|13622288|gb|AAK34023.1|
gi|13622289|gb|AAK34024.1|
gi|13622290|gb|AAK34025.1|
gi|13622294|gb|AAK34029.1|
gi|13622295|gb|AAK34030.1|
gi|13622296|gb|AAK34031.1|
gi|13622297|gb|AAK34032.1|
gi|13622298|gb|AAK34033.1|
gi|13622299|gb|AAK34034.1|
gi|13622301|gb|AAK34035.1|
gi|13622306|gb|AAK34040.1|
gi|13622326|gb|AAK34058.1|
gi|13622328|gb|AAK34060.1|
gi|13622329|gb|AAK34061.1|
gi|13622330|gb|AAK34062.1|
gi|13622332|gb|AAK34064.1|
gi|13622333|gb|AAK34065.1|
gi|13622335|gb|AAK34066.1|
gi|13622338|gb|AAK34069.1|
gi|13622339|gb|AAK34070.1|
gi|13622340|gb|AAK34071.1|
gi|13622341|gb|AAK34072.1|
gi|13622343|gb|AAK34073.1|
gi|13622350|gb|AAK34080.1|
gi|13622351|gb|AAK34081.1|
gi|13622352|gb|AAK34082.1|
gi|13622353|gb|AAK34083.1|
gi|13622355|gb|AAK34084.1|
gi|13622356|gb|AAK34085.1|
gi|13622357|gb|AAK34086.1|
gi|13622358|gb|AAK34087.1|
gi|13622359|gb|AAK34088.1|
gi|13622360|gb|AAK34089.1|
gi|13622361|gb|AAK34090.1|
gi|13622362|gb|AAK34091.1|
gi|13622363|gb|AAK34092.1|
gi|13622364|gb|AAK34093.1|
gi|13622366|gb|AAK34094.1|
gi|13622367|gb|AAK34095.1|
gi|13622368|gb|AAK34096.1|

Table 33: List of GAS ORFs which are shared with GBS and Spn

gi|13622369|gb|AAK34097.1|
gi|13622370|gb|AAK34098.1|
gi|13622371|gb|AAK34099.1|
gi|13622372|gb|AAK34100.1|
gi|13622373|gb|AAK34101.1|
gi|13622374|gb|AAK34102.1|
gi|13622375|gb|AAK34103.1|
gi|13622376|gb|AAK34104.1|
gi|13622377|gb|AAK34105.1|
gi|13622378|gb|AAK34106.1|
gi|13622380|gb|AAK34107.1|
gi|13622383|gb|AAK34110.1|
gi|13622384|gb|AAK34111.1|
gi|13622387|gb|AAK34114.1|
gi|13622389|gb|AAK34116.1|
gi|13622394|gb|AAK34120.1|
gi|13622395|gb|AAK34121.1|
gi|13622396|gb|AAK34122.1|
gi|13622398|gb|AAK34124.1|
gi|13622399|gb|AAK34125.1|
gi|13622400|gb|AAK34126.1|
gi|13622401|gb|AAK34127.1|
gi|13622403|gb|AAK34128.1|
gi|13622405|gb|AAK34130.1|
gi|13622406|gb|AAK34131.1|
gi|13622407|gb|AAK34132.1|
gi|13622408|gb|AAK34133.1|
gi|13622415|gb|AAK34139.1|
gi|13622416|gb|AAK34140.1|
gi|13622417|gb|AAK34141.1|
gi|13622419|gb|AAK34143.1|
gi|13622420|gb|AAK34144.1|
gi|13622424|gb|AAK34147.1|
gi|13622425|gb|AAK34148.1|
gi|13622431|gb|AAK34153.1|
gi|13622432|gb|AAK34154.1|
gi|13622433|gb|AAK34155.1|
gi|13622434|gb|AAK34156.1|
gi|13622435|gb|AAK34157.1|
gi|13622436|gb|AAK34158.1|
gi|13622437|gb|AAK34159.1|
gi|13622444|gb|AAK34165.1|
gi|13622447|gb|AAK34168.1|
gi|13622450|gb|AAK34170.1|
gi|13622451|gb|AAK34171.1|
gi|13622455|gb|AAK34175.1|
gi|13622457|gb|AAK34177.1|
gi|13622458|gb|AAK34178.1|
gi|13622460|gb|AAK34179.1|
gi|13622461|gb|AAK34180.1|
gi|13622462|gb|AAK34181.1|
gi|13622463|gb|AAK34182.1|
gi|13622464|gb|AAK34183.1|
gi|13622465|gb|AAK34184.1|
gi|13622467|gb|AAK34186.1|
gi|13622468|gb|AAK34187.1|
gi|13622471|gb|AAK34189.1|
gi|13622473|gb|AAK34191.1|
gi|13622474|gb|AAK34192.1|
gi|13622477|gb|AAK34195.1|
gi|13622478|gb|AAK34196.1|
gi|13622479|gb|AAK34197.1|
gi|13622481|gb|AAK34198.1|
gi|13622482|gb|AAK34199.1|
gi|13622483|gb|AAK34200.1|
gi|13622484|gb|AAK34201.1|
gi|13622485|gb|AAK34202.1|
gi|13622486|gb|AAK34203.1|
gi|13622491|gb|AAK34207.1|
gi|13622492|gb|AAK34208.1|
gi|13622493|gb|AAK34209.1|
gi|13622494|gb|AAK34210.1|
gi|13622495|gb|AAK34211.1|
gi|13622496|gb|AAK34212.1|
gi|13622497|gb|AAK34213.1|
gi|13622499|gb|AAK34214.1|
gi|13622500|gb|AAK34215.1|
gi|13622501|gb|AAK34216.1|
gi|13622506|gb|AAK34221.1|
gi|13622507|gb|AAK34222.1|
gi|13622508|gb|AAK34223.1|
gi|13622509|gb|AAK34224.1|
gi|13622511|gb|AAK34225.1|
gi|13622512|gb|AAK34226.1|
gi|13622513|gb|AAK34227.1|
gi|13622515|gb|AAK34229.1|
gi|13622516|gb|AAK34230.1|
gi|13622517|gb|AAK34231.1|
gi|13622518|gb|AAK34232.1|
gi|13622520|gb|AAK34233.1|
gi|13622521|gb|AAK34234.1|
gi|13622523|gb|AAK34236.1|
gi|13622524|gb|AAK34237.1|
gi|13622525|gb|AAK34238.1|
gi|13622526|gb|AAK34239.1|
gi|13622527|gb|AAK34240.1|
gi|13622579|gb|AAK34289.1|
gi|13622583|gb|AAK34292.1|
gi|13622585|gb|AAK34294.1|
gi|13622587|gb|AAK34296.1|
gi|13622588|gb|AAK34297.1|
gi|13622590|gb|AAK34299.1|
gi|13622591|gb|AAK34300.1|
gi|13622593|gb|AAK34301.1|
gi|13622595|gb|AAK34303.1|
gi|13622596|gb|AAK34304.1|
gi|13622597|gb|AAK34305.1|
gi|13622598|gb|AAK34306.1|
gi|13622599|gb|AAK34307.1|
gi|13622600|gb|AAK34308.1|
gi|13622601|gb|AAK34309.1|
gi|13622603|gb|AAK34310.1|

Table 33: List of GAS ORFs which are shared with GBS and Spn

gi|13622604|gb|AAK34311.1|
gi|13622606|gb|AAK34313.1|
gi|13622607|gb|AAK34314.1|
gi|13622608|gb|AAK34315.1|
gi|13622609|gb|AAK34316.1|
gi|13622610|gb|AAK34317.1|
gi|13622611|gb|AAK34318.1|
gi|13622612|gb|AAK34319.1|
gi|13622615|gb|AAK34321.1|
gi|13622616|gb|AAK34322.1|
gi|13622617|gb|AAK34323.1|
gi|13622618|gb|AAK34324.1|
gi|13622621|gb|AAK34327.1|
gi|13622622|gb|AAK34328.1|
gi|13622623|gb|AAK34329.1|
gi|13622624|gb|AAK34330.1|
gi|13622625|gb|AAK34331.1|
gi|13622626|gb|AAK34332.1|
gi|13622628|gb|AAK34333.1|
gi|13622629|gb|AAK34334.1|
gi|13622630|gb|AAK34335.1|
gi|13622631|gb|AAK34336.1|
gi|13622632|gb|AAK34337.1|
gi|13622634|gb|AAK34339.1|
gi|13622636|gb|AAK34341.1|
gi|13622640|gb|AAK34344.1|
gi|13622641|gb|AAK34345.1|
gi|13622652|gb|AAK34355.1|
gi|13622653|gb|AAK34356.1|
gi|13622654|gb|AAK34357.1|
gi|13622656|gb|AAK34359.1|
gi|13622660|gb|AAK34363.1|
gi|13622665|gb|AAK34367.1|
gi|13622668|gb|AAK34370.1|
gi|13622675|gb|AAK34376.1|
gi|13622676|gb|AAK34377.1|
gi|13622683|gb|AAK34383.1|
gi|13622684|gb|AAK34384.1|
gi|13622685|gb|AAK34385.1|
gi|13622688|gb|AAK34387.1|
gi|13622689|gb|AAK34388.1|
gi|13622690|gb|AAK34389.1|
gi|13622691|gb|AAK34390.1|
gi|13622692|gb|AAK34391.1|
gi|13622693|gb|AAK34392.1|
gi|13622694|gb|AAK34393.1|
gi|13622695|gb|AAK34394.1|
gi|13622696|gb|AAK34395.1|
gi|13622698|gb|AAK34396.1|
gi|13622699|gb|AAK34397.1|
gi|13622700|gb|AAK34398.1|
gi|13622701|gb|AAK34399.1|
gi|13622702|gb|AAK34400.1|
gi|13622703|gb|AAK34401.1|
gi|13622704|gb|AAK34402.1|
gi|13622705|gb|AAK34403.1|
gi|13622711|gb|AAK34408.1|
gi|13622713|gb|AAK34410.1|
gi|13622714|gb|AAK34411.1|
gi|13622715|gb|AAK34412.1|
gi|13622718|gb|AAK34414.1|
gi|13622719|gb|AAK34415.1|
gi|13622720|gb|AAK34416.1|
gi|13622721|gb|AAK34417.1|
gi|13622722|gb|AAK34418.1|
gi|13622723|gb|AAK34419.1|
gi|13622727|gb|AAK34422.1|
gi|13622728|gb|AAK34423.1|
gi|13622729|gb|AAK34424.1|
gi|13622730|gb|AAK34425.1|
gi|13622731|gb|AAK34426.1|
gi|13622733|gb|AAK34428.1|
gi|13622734|gb|AAK34429.1|
gi|13622735|gb|AAK34430.1|
gi|13622736|gb|AAK34431.1|
gi|13622737|gb|AAK34432.1|
gi|13622740|gb|AAK34434.1|
gi|13622741|gb|AAK34435.1|
gi|13622742|gb|AAK34436.1|
gi|13622744|gb|AAK34438.1|
gi|13622745|gb|AAK34439.1|
gi|13622746|gb|AAK34440.1|
gi|13622749|gb|AAK34442.1|
gi|13622750|gb|AAK34443.1|
gi|13622751|gb|AAK34444.1|
gi|13622752|gb|AAK34445.1|
gi|13622753|gb|AAK34446.1|
gi|13622754|gb|AAK34447.1|
gi|13622760|gb|AAK34452.1|
gi|13622762|gb|AAK34454.1|
gi|13622763|gb|AAK34455.1|
gi|13622764|gb|AAK34456.1|
gi|13622765|gb|AAK34457.1|
gi|13622766|gb|AAK34458.1|
gi|13622767|gb|AAK34459.1|
gi|13622768|gb|AAK34460.1|
gi|13622770|gb|AAK34462.1|
gi|13622771|gb|AAK34463.1|
gi|13622774|gb|AAK34465.1|
gi|13622775|gb|AAK34466.1|
gi|13622776|gb|AAK34467.1|
gi|13622777|gb|AAK34468.1|
gi|13622778|gb|AAK34469.1|
gi|13622779|gb|AAK34470.1|
gi|13622780|gb|AAK34471.1|
gi|13622781|gb|AAK34472.1|
gi|13622782|gb|AAK34473.1|
gi|13622783|gb|AAK34474.1|
gi|13622785|gb|AAK34475.1|
gi|13622787|gb|AAK34477.1|
gi|13622789|gb|AAK34479.1|
gi|13622790|gb|AAK34480.1|

Table 33: List of GAS ORFs which are shared with GBS and Spn

gi|13622791|gb|AAK34481.1|
gi|13622792|gb|AAK34482.1|
gi|13622793|gb|AAK34483.1|
gi|13622794|gb|AAK34484.1|
gi|13622795|gb|AAK34485.1|
gi|13622796|gb|AAK34486.1|
gi|13622798|gb|AAK34487.1|
gi|13622799|gb|AAK34488.1|
gi|13622800|gb|AAK34489.1|
gi|13622801|gb|AAK34490.1|
gi|13622802|gb|AAK34491.1|
gi|13622803|gb|AAK34492.1|
gi|13622804|gb|AAK34493.1|
gi|13622805|gb|AAK34494.1|
gi|13622806|gb|AAK34495.1|
gi|13622807|gb|AAK34496.1|
gi|13622808|gb|AAK34497.1|
gi|13622809|gb|AAK34498.1|
gi|13622810|gb|AAK34499.1|
gi|13622812|gb|AAK34500.1|
gi|13622813|gb|AAK34501.1|
gi|13622814|gb|AAK34502.1|
gi|13622815|gb|AAK34503.1|
gi|13622818|gb|AAK34506.1|
gi|13622821|gb|AAK34509.1|
gi|13622822|gb|AAK34510.1|
gi|13622823|gb|AAK34511.1|
gi|13622825|gb|AAK34512.1|
gi|13622826|gb|AAK34513.1|
gi|13622827|gb|AAK34514.1|
gi|13622828|gb|AAK34515.1|
gi|13622829|gb|AAK34516.1|
gi|13622830|gb|AAK34517.1|
gi|13622833|gb|AAK34520.1|
gi|13622838|gb|AAK34524.1|
gi|13622839|gb|AAK34525.1|
gi|13622840|gb|AAK34526.1|
gi|13622841|gb|AAK34527.1|
gi|13622847|gb|AAK34532.1|
gi|13622848|gb|AAK34533.1|
gi|13622849|gb|AAK34534.1|
gi|13622853|gb|AAK34537.1|
gi|13622854|gb|AAK34538.1|
gi|13622856|gb|AAK34540.1|
gi|13622857|gb|AAK34541.1|
gi|13622858|gb|AAK34542.1|
gi|13622860|gb|AAK34543.1|
gi|13622861|gb|AAK34544.1|
gi|13622862|gb|AAK34545.1|
gi|13622863|gb|AAK34546.1|
gi|13622864|gb|AAK34547.1|
gi|13622865|gb|AAK34548.1|
gi|13622866|gb|AAK34549.1|
gi|13622867|gb|AAK34550.1|
gi|13622868|gb|AAK34551.1|
gi|13622869|gb|AAK34552.1|
gi|13622870|gb|AAK34553.1|
gi|13622873|gb|AAK34555.1|
gi|13622875|gb|AAK34557.1|
gi|13622876|gb|AAK34558.1|
gi|13622877|gb|AAK34559.1|
gi|13622878|gb|AAK34560.1|
gi|13622879|gb|AAK34561.1|
gi|13622880|gb|AAK34562.1|
gi|13622881|gb|AAK34563.1|
gi|13622882|gb|AAK34564.1|
gi|13622885|gb|AAK34566.1|
gi|13622886|gb|AAK34567.1|
gi|13622887|gb|AAK34568.1|
gi|13622888|gb|AAK34569.1|
gi|13622890|gb|AAK34571.1|
gi|13622893|gb|AAK34574.1|
gi|13622896|gb|AAK34576.1|
gi|13622898|gb|AAK34578.1|
gi|13622899|gb|AAK34579.1|
gi|13622900|gb|AAK34580.1|
gi|13622901|gb|AAK34581.1|
gi|13622903|gb|AAK34583.1|
gi|13622905|gb|AAK34585.1|
gi|13622906|gb|AAK34586.1|
gi|13622907|gb|AAK34587.1|
gi|13622908|gb|AAK34588.1|
gi|13622910|gb|AAK34589.1|
gi|13622911|gb|AAK34590.1|
gi|13622912|gb|AAK34591.1|
gi|13622913|gb|AAK34592.1|
gi|13622914|gb|AAK34593.1|
gi|13622915|gb|AAK34594.1|
gi|13622917|gb|AAK34596.1|
gi|13622918|gb|AAK34597.1|
gi|13622919|gb|AAK34598.1|
gi|13622921|gb|AAK34599.1|
gi|13622922|gb|AAK34600.1|
gi|13622924|gb|AAK34602.1|
gi|13622925|gb|AAK34603.1|
gi|13622926|gb|AAK34604.1|
gi|13622927|gb|AAK34605.1|
gi|13622928|gb|AAK34606.1|
gi|13622929|gb|AAK34607.1|
gi|13622930|gb|AAK34608.1|
gi|13622931|gb|AAK34609.1|
gi|13622933|gb|AAK34610.1|
gi|13622941|gb|AAK34617.1|
gi|13622944|gb|AAK34620.1|
gi|13622945|gb|AAK34621.1|
gi|13622947|gb|AAK34623.1|
gi|13622948|gb|AAK34624.1|
gi|13622949|gb|AAK34625.1|
gi|13622950|gb|AAK34626.1|
gi|13622952|gb|AAK34627.1|
gi|13622955|gb|AAK34630.1|
gi|13622956|gb|AAK34631.1|

Table 33: List of GAS ORFs which are shared with GBS and Spn

gi|13622959|gb|AAK34634.1|
gi|13622961|gb|AAK34636.1|
gi|13622963|gb|AAK34638.1|
gi|13622964|gb|AAK34639.1|
gi|13622967|gb|AAK34641.1|
gi|13622969|gb|AAK34643.1|
gi|13622971|gb|AAK34645.1|
gi|13622973|gb|AAK34647.1|
gi|13622974|gb|AAK34648.1|
gi|13622977|gb|AAK34651.1|
gi|13622981|gb|AAK34654.1|
gi|13622982|gb|AAK34655.1|
gi|13622983|gb|AAK34656.1|
gi|13622984|gb|AAK34657.1|
gi|13622985|gb|AAK34658.1|
gi|13622989|gb|AAK34661.1|
gi|13622990|gb|AAK34662.1|
gi|13622991|gb|AAK34663.1|
gi|13622992|gb|AAK34664.1|
gi|13622995|gb|AAK34666.1|
gi|13622996|gb|AAK34667.1|
gi|13622998|gb|AAK34669.1|
gi|13622999|gb|AAK34670.1|
gi|13623000|gb|AAK34671.1|
gi|13623001|gb|AAK34672.1|
gi|13623002|gb|AAK34673.1|
gi|13623004|gb|AAK34674.1|
gi|13623005|gb|AAK34675.1|
gi|13623006|gb|AAK34676.1|
gi|13623007|gb|AAK34677.1|
gi|13623009|gb|AAK34679.1|
gi|13623019|gb|AAK34688.1|
gi|13623020|gb|AAK34689.1|
gi|13623030|gb|AAK34698.1|
gi|13623031|gb|AAK34699.1|
gi|13623032|gb|AAK34700.1|
gi|13623033|gb|AAK34701.1|
gi|13623038|gb|AAK34705.1|
gi|13623045|gb|AAK34712.1|
gi|13623046|gb|AAK34713.1|
gi|13623047|gb|AAK34714.1|
gi|13623049|gb|AAK34715.1|
gi|13623050|gb|AAK34716.1|
gi|13623051|gb|AAK34717.1|
gi|13623052|gb|AAK34718.1|
gi|13623053|gb|AAK34719.1|
gi|13623054|gb|AAK34720.1|
gi|13623056|gb|AAK34722.1|
gi|13623058|gb|AAK34724.1|
gi|13623062|gb|AAK34727.1|
gi|13623064|gb|AAK34729.1|
gi|13623065|gb|AAK34730.1|
gi|13623069|gb|AAK34733.1|
gi|13623074|gb|AAK34738.1|
gi|13623081|gb|AAK34744.1|
gi|13623082|gb|AAK34745.1|
gi|13623083|gb|AAK34746.1|
gi|13623085|gb|AAK34747.1|
gi|13623086|gb|AAK34748.1|
gi|13623088|gb|AAK34750.1|
gi|13623089|gb|AAK34751.1|
gi|13623090|gb|AAK34752.1|
gi|13623091|gb|AAK34753.1|
gi|13623093|gb|AAK34755.1|
gi|13623095|gb|AAK34756.1|
gi|13623096|gb|AAK34757.1|
gi|13623098|gb|AAK34759.1|
gi|13623099|gb|AAK34760.1|
gi|13623100|gb|AAK34761.1|
gi|13623102|gb|AAK34763.1|
gi|13623103|gb|AAK34764.1|
gi|13623105|gb|AAK34766.1|
gi|13623107|gb|AAK34767.1|
gi|13623128|gb|AAK34787.1|
gi|13623129|gb|AAK34788.1|
gi|13623131|gb|AAK34790.1|
gi|13623132|gb|AAK34791.1|
gi|13623133|gb|AAK34792.1|
gi|13623134|gb|AAK34793.1|
gi|13623136|gb|AAK34794.1|
gi|13623138|gb|AAK34796.1|
gi|13623139|gb|AAK34797.1|
gi|13623150|gb|AAK34807.1|
gi|13623151|gb|AAK34808.1|
gi|13623152|gb|AAK34809.1|
gi|13623154|gb|AAK34811.1|
gi|13623155|gb|AAK34812.1|
gi|13623156|gb|AAK34813.1|
gi|13623157|gb|AAK34814.1|
gi|13623159|gb|AAK34815.1|
gi|13623161|gb|AAK34817.1|
gi|13623162|gb|AAK34818.1|
gi|13623163|gb|AAK34819.1|
gi|13623165|gb|AAK34821.1|
gi|13623166|gb|AAK34822.1|
gi|13623167|gb|AAK34823.1|
gi|13623168|gb|AAK34824.1|
gi|13623170|gb|AAK34826.1|
gi|13623171|gb|AAK34827.1|
gi|13623175|gb|AAK34830.1|
gi|13623176|gb|AAK34831.1|
gi|13623177|gb|AAK34832.1|
gi|13623179|gb|AAK34834.1|
gi|13623180|gb|AAK34835.1|
gi|13623182|gb|AAK34836.1|
gi|13623183|gb|AAK34837.1|
gi|13623184|gb|AAK34838.1|
gi|13623185|gb|AAK34839.1|
gi|13623186|gb|AAK34840.1|
gi|13623187|gb|AAK34841.1|

Table 34: List of GAS ORF's which are shared with GBS but not with Spn

gi|13621381|gb|AAK33195.1|
gi|13621423|gb|AAK33233.1|
gi|13621440|gb|AAK33249.1|
gi|13621443|gb|AAK33251.1|
gi|13621453|gb|AAK33260.1|
gi|13621454|gb|AAK33261.1|
gi|13621479|gb|AAK33284.1|
gi|13621482|gb|AAK33287.1|
gi|13621492|gb|AAK33296.1|
gi|13621493|gb|AAK33297.1|
gi|13621497|gb|AAK33300.1|
gi|13621498|gb|AAK33301.1|
gi|13621512|gb|AAK33314.1|
gi|13621514|gb|AAK33316.1|
gi|13621556|gb|AAK33354.1|
gi|13621570|gb|AAK33366.1|
gi|13621587|gb|AAK33382.1|
gi|13621610|gb|AAK33403.1|
gi|13621613|gb|AAK33405.1|
gi|13621626|gb|AAK33418.1|
gi|13621632|gb|AAK33423.1|
gi|13621635|gb|AAK33426.1|
gi|13621643|gb|AAK33433.1|
gi|13621655|gb|AAK33444.1|
gi|13621656|gb|AAK33445.1|
gi|13621659|gb|AAK33448.1|
gi|13621673|gb|AAK33461.1|
gi|13621686|gb|AAK33473.1|
gi|13621696|gb|AAK33482.1|
gi|13621703|gb|AAK33488.1|
gi|13621712|gb|AAK33497.1|
gi|13621728|gb|AAK33511.1|
gi|13621738|gb|AAK33520.1|
gi|13621739|gb|AAK33521.1|
gi|13621740|gb|AAK33522.1|
gi|13621772|gb|AAK33551.1|
gi|13621776|gb|AAK33555.1|
gi|13621791|gb|AAK33569.1|
gi|13621798|gb|AAK33575.1|
gi|13621801|gb|AAK33578.1|
gi|13621803|gb|AAK33580.1|
gi|13621804|gb|AAK33581.1|
gi|13621832|gb|AAK33606.1|
gi|13621833|gb|AAK33607.1|
gi|13621896|gb|AAK33665.1|
gi|13621897|gb|AAK33666.1|
gi|13621906|gb|AAK33674.1|
gi|13621911|gb|AAK33679.1|
gi|13621949|gb|AAK33713.1|
gi|13621951|gb|AAK33715.1|
gi|13621962|gb|AAK33724.1|
gi|13621963|gb|AAK33725.1|
gi|13621964|gb|AAK33726.1|
gi|13621971|gb|AAK33732.1|
gi|13621976|gb|AAK33737.1|
gi|13621983|gb|AAK33744.1|
gi|13621988|gb|AAK33748.1|
gi|13622014|gb|AAK33772.1|
gi|13622015|gb|AAK33773.1|
gi|13622022|gb|AAK33779.1|
gi|13622023|gb|AAK33780.1|
gi|13622028|gb|AAK33784.1|
gi|13622029|gb|AAK33785.1|
gi|13622037|gb|AAK33792.1|
gi|13622038|gb|AAK33793.1|
gi|13622040|gb|AAK33795.1|
gi|13622057|gb|AAK33811.1|
gi|13622061|gb|AAK33814.1|
gi|13622063|gb|AAK33816.1|
gi|13622066|gb|AAK33819.1|
gi|13622067|gb|AAK33820.1|
gi|13622076|gb|AAK33828.1|
gi|13622078|gb|AAK33830.1|
gi|13622084|gb|AAK33835.1|
gi|13622098|gb|AAK33848.1|
gi|13622099|gb|AAK33849.1|
gi|13622100|gb|AAK33850.1|
gi|13622104|gb|AAK33854.1|
gi|13622110|gb|AAK33859.1|
gi|13622116|gb|AAK33865.1|
gi|13622124|gb|AAK33873.1|
gi|13622159|gb|AAK33905.1|
gi|13622193|gb|AAK33936.1|
gi|13622194|gb|AAK33937.1|
gi|13622195|gb|AAK33938.1|
gi|13622196|gb|AAK33939.1|
gi|13622202|gb|AAK33944.1|
gi|13622203|gb|AAK33945.1|
gi|13622206|gb|AAK33948.1|
gi|13622210|gb|AAK33951.1|
gi|13622221|gb|AAK33961.1|
gi|13622231|gb|AAK33971.1|
gi|13622234|gb|AAK33973.1|
gi|13622238|gb|AAK33977.1|
gi|13622245|gb|AAK33984.1|
gi|13622246|gb|AAK33985.1|
gi|13622248|gb|AAK33986.1|
gi|13622249|gb|AAK33987.1|
gi|13622251|gb|AAK33989.1|
gi|13622254|gb|AAK33992.1|
gi|13622267|gb|AAK34004.1|
gi|13622291|gb|AAK34026.1|
gi|13622302|gb|AAK34036.1|
gi|13622303|gb|AAK34037.1|
gi|13622304|gb|AAK34038.1|
gi|13622327|gb|AAK34059.1|
gi|13622344|gb|AAK34074.1|
gi|13622345|gb|AAK34075.1|
gi|13622346|gb|AAK34076.1|
gi|13622347|gb|AAK34077.1|
gi|13622348|gb|AAK34078.1|
gi|13622349|gb|AAK34079.1|

Table 34: List of GAS ORF's which are shared with GBS but not with Spn

gi|13622382|gb|AAK34109.1|
gi|13622386|gb|AAK34113.1|
gi|13622391|gb|AAK34118.1|
gi|13622392|gb|AAK34119.1|
gi|13622397|gb|AAK34123.1|
gi|13622404|gb|AAK34129.1|
gi|13622412|gb|AAK34136.1|
gi|13622413|gb|AAK34137.1|
gi|13622414|gb|AAK34138.1|
gi|13622418|gb|AAK34142.1|
gi|13622430|gb|AAK34152.1|
gi|13622446|gb|AAK34167.1|
gi|13622449|gb|AAK34169.1|
gi|13622453|gb|AAK34173.1|
gi|13622470|gb|AAK34188.1|
gi|13622487|gb|AAK34204.1|
gi|13622490|gb|AAK34206.1|
gi|13622502|gb|AAK34217.1|
gi|13622503|gb|AAK34218.1|
gi|13622514|gb|AAK34228.1|
gi|13622528|gb|AAK34241.1|
gi|13622540|gb|AAK34252.1|
gi|13622541|gb|AAK34253.1|
gi|13622544|gb|AAK34255.1|
gi|13622545|gb|AAK34256.1|
gi|13622546|gb|AAK34257.1|
gi|13622547|gb|AAK34258.1|
gi|13622548|gb|AAK34259.1|
gi|13622550|gb|AAK34261.1|
gi|13622551|gb|AAK34262.1|
gi|13622552|gb|AAK34263.1|
gi|13622556|gb|AAK34267.1|
gi|13622557|gb|AAK34268.1|
gi|13622558|gb|AAK34269.1|
gi|13622559|gb|AAK34270.1|
gi|13622563|gb|AAK34273.1|
gi|13622571|gb|AAK34281.1|
gi|13622576|gb|AAK34286.1|
gi|13622581|gb|AAK34290.1|
gi|13622582|gb|AAK34291.1|
gi|13622586|gb|AAK34295.1|
gi|13622589|gb|AAK34298.1|
gi|13622605|gb|AAK34312.1|
gi|13622633|gb|AAK34338.1|
gi|13622635|gb|AAK34340.1|
gi|13622637|gb|AAK34342.1|
gi|13622638|gb|AAK34343.1|
gi|13622657|gb|AAK34360.1|
gi|13622707|gb|AAK34404.1|
gi|13622716|gb|AAK34413.1|
gi|13622724|gb|AAK34420.1|
gi|13622732|gb|AAK34427.1|
gi|13622743|gb|AAK34437.1|
gi|13622761|gb|AAK34453.1|
gi|13622773|gb|AAK34464.1|
gi|13622788|gb|AAK34478.1|
gi|13622816|gb|AAK34504.1|
gi|13622817|gb|AAK34505.1|
gi|13622846|gb|AAK34531.1|
gi|13622852|gb|AAK34536.1|
gi|13622874|gb|AAK34556.1|
gi|13622889|gb|AAK34570.1|
gi|13622891|gb|AAK34572.1|
gi|13622892|gb|AAK34573.1|
gi|13622897|gb|AAK34577.1|
gi|13622902|gb|AAK34582.1|
gi|13622904|gb|AAK34584.1|
gi|13622916|gb|AAK34595.1|
gi|13622923|gb|AAK34601.1|
gi|13622934|gb|AAK34611.1|
gi|13622953|gb|AAK34628.1|
gi|13622954|gb|AAK34629.1|
gi|13622960|gb|AAK34635.1|
gi|13622968|gb|AAK34642.1|
gi|13622980|gb|AAK34653.1|
gi|13622987|gb|AAK34659.1|
gi|13623012|gb|AAK34682.1|
gi|13623013|gb|AAK34683.1|
gi|13623014|gb|AAK34684.1|
gi|13623015|gb|AAK34685.1|
gi|13623016|gb|AAK34686.1|
gi|13623018|gb|AAK34687.1|
gi|13623022|gb|AAK34691.1|
gi|13623029|gb|AAK34697.1|
gi|13623037|gb|AAK34704.1|
gi|13623055|gb|AAK34721.1|
gi|13623060|gb|AAK34725.1|
gi|13623061|gb|AAK34726.1|
gi|13623063|gb|AAK34728.1|
gi|13623066|gb|AAK34731.1|
gi|13623068|gb|AAK34732.1|
gi|13623092|gb|AAK34754.1|
gi|13623097|gb|AAK34758.1|
gi|13623104|gb|AAK34765.1|
gi|13623126|gb|AAK34785.1|
gi|13623130|gb|AAK34789.1|
gi|13623137|gb|AAK34795.1|
gi|13623153|gb|AAK34810.1|
gi|13623164|gb|AAK34820.1|
gi|13623178|gb|AAK34833.1|

Table 35: GAS ORF's which are shared with pneumococcus but not with GBS

gi|13621338|gb|AAK33157.1|
gi|13621352|gb|AAK33168.1|
gi|13621410|gb|AAK33221.1|
gi|13621433|gb|AAK33242.1|
gi|13621445|gb|AAK33253.1|
gi|13621446|gb|AAK33254.1|
gi|13621447|gb|AAK33255.1|
gi|13621448|gb|AAK33256.1|
gi|13621449|gb|AAK33257.1|
gi|13621451|gb|AAK33259.1|
gi|13621460|gb|AAK33267.1|
gi|13621466|gb|AAK33272.1|
gi|13621489|gb|AAK33293.1|
gi|13621490|gb|AAK33294.1|
gi|13621519|gb|AAK33320.1|
gi|13621520|gb|AAK33321.1|
gi|13621653|gb|AAK33443.1|
gi|13621722|gb|AAK33506.1|
gi|13621723|gb|AAK33507.1|
gi|13621724|gb|AAK33508.1|
gi|13621805|gb|AAK33582.1|
gi|13621900|gb|AAK33669.1|
gi|13622011|gb|AAK33769.1|
gi|13622212|gb|AAK33953.1|
gi|13622280|gb|AAK34016.1|
gi|13622381|gb|AAK34108.1|
gi|13622409|gb|AAK34134.1|
gi|13622410|gb|AAK34135.1|
gi|13622423|gb|AAK34146.1|
gi|13622428|gb|AAK34151.1|
gi|13622441|gb|AAK34162.1|
gi|13622442|gb|AAK34163.1|
gi|13622454|gb|AAK34174.1|
gi|13622456|gb|AAK34176.1|
gi|13622619|gb|AAK34325.1|
gi|13622642|gb|AAK34346.1|
gi|13622643|gb|AAK34347.1|
gi|13622664|gb|AAK34366.1|
gi|13622666|gb|AAK34368.1|
gi|13622667|gb|AAK34369.1|
gi|13622671|gb|AAK34372.1|
gi|13622672|gb|AAK34373.1|
gi|13622673|gb|AAK34374.1|
gi|13622674|gb|AAK34375.1|
gi|13622679|gb|AAK34380.1|
gi|13622680|gb|AAK34381.1|
gi|13622682|gb|AAK34382.1|
gi|13622755|gb|AAK34448.1|
gi|13622758|gb|AAK34450.1|
gi|13622759|gb|AAK34451.1|
gi|13622835|gb|AAK34521.1|
gi|13622837|gb|AAK34523.1|
gi|13622937|gb|AAK34614.1|
gi|13622942|gb|AAK34618.1|
gi|13622946|gb|AAK34622.1|
gi|13622978|gb|AAK34652.1|
gi|13623027|gb|AAK34695.1|
gi|13623087|gb|AAK34749.1|
gi|13623101|gb|AAK34762.1|
gi|13623144|gb|AAK34802.1|
gi|13623146|gb|AAK34804.1|
gi|13623147|gb|AAK34805.1|

Table 36: Spn ORF's are shared with GBS and GAS

SP0001
SP0002
SP0003
SP0004
SP0005
SP0006
SP0007
SP0008
SP0010
SP0011
SP0013
SP0014
SP0019
SP0021
SP0024
SP0027
SP0032
SP0033
SP0034
SP0035
SP0036
SP0037
SP0042
SP0044
SP0045
SP0046
SP0047
SP0048
SP0051
SP0053
SP0054
SP0056
SP0063
SP0073
SP0074
SP0078
SP0079
SP0083
SP0084
SP0085
SP0095
SP0105
SP0106
SP0111
SP0112
SP0118
SP0120
SP0121
SP0122
SP0127
SP0128
SP0129
SP0148
SP0149
SP0151
SP0152
SP0158
SP0173
SP0179
SP0180
SP0184
SP0185
SP0186
SP0187
SP0189
SP0192
SP0194
SP0197
SP0199
SP0202
SP0204
SP0205
SP0208
SP0209
SP0210
SP0211
SP0212
SP0213
SP0214
SP0215
SP0216
SP0217
SP0218
SP0219
SP0220
SP0221
SP0222
SP0224
SP0225
SP0226
SP0227
SP0228
SP0229
SP0230
SP0231
SP0232
SP0233
SP0234
SP0235
SP0236
SP0240
SP0242
SP0243
SP0245
SP0246
SP0247
SP0248
SP0249
SP0250
SP0251
SP0252
SP0253
SP0254
SP0259
SP0261
SP0262
SP0263
SP0264
SP0265
SP0266
SP0268
SP0271
SP0272
SP0273
SP0274
SP0280
SP0281
SP0282
SP0283
SP0284
SP0285
SP0286
SP0287
SP0289
SP0290
SP0291
SP0292
SP0294
SP0295
SP0303
SP0310
SP0314
SP0317
SP0318
SP0319
SP0320
SP0321
SP0322
SP0323
SP0324
SP0325
SP0327
SP0330
SP0334
SP0336
SP0337
SP0338
SP0340
SP0342
SP0369
SP0370
SP0371
SP0373
SP0374
SP0381
SP0382
SP0383
SP0384
SP0385
SP0386
SP0387
SP0400
SP0401
SP0402
SP0403
SP0404
SP0405
SP0406
SP0408
SP0410
SP0411
SP0412
SP0415
SP0416
SP0417
SP0418
SP0419
SP0420
SP0421
SP0422
SP0423
SP0424
SP0425
SP0426
SP0427
SP0433
SP0434
SP0435
SP0436
SP0437
SP0438
SP0439
SP0441
SP0442
SP0443
SP0452
SP0453
SP0454
SP0457
SP0458
SP0459
SP0461
SP0466
SP0467
SP0474
SP0477
SP0478
SP0483
SP0486
SP0488
SP0489
SP0493
SP0494
SP0499

Table 36: Spn ORF's are shared with GBS and GAS

SP0500
SP0501
SP0502
SP0515
SP0516
SP0517
SP0519
SP0521
SP0522
SP0523
SP0526
SP0549
SP0550
SP0552
SP0553
SP0554
SP0555
SP0556
SP0557
SP0563
SP0567
SP0568
SP0576
SP0577
SP0578
SP0579
SP0581
SP0588
SP0589
SP0591
SP0592
SP0593
SP0603
SP0604
SP0605
SP0608
SP0610
SP0611
SP0613
SP0614
SP0615
SP0616
SP0618
SP0620
SP0622
SP0623
SP0624
SP0626
SP0630
SP0631
SP0636
SP0637
SP0638
SP0645
SP0646
SP0647
SP0652
SP0657
SP0660
SP0662
SP0663
SP0665
SP0668
SP0669
SP0671
SP0672
SP0673
SP0674
SP0675
SP0676
SP0678
SP0680
SP0681
SP0687
SP0688
SP0689
SP0690
SP0701
SP0702
SP0709
SP0713
SP0726
SP0727
SP0729
SP0735
SP0736
SP0741
SP0744
SP0745
SP0746
SP0756
SP0757
SP0758
SP0760
SP0761
SP0762
SP0764
SP0765
SP0766
SP0767
SP0768
SP0770
SP0771
SP0775
SP0776
SP0778
SP0779
SP0780
SP0782
SP0784
SP0785
SP0786
SP0787
SP0788
SP0792
SP0793
SP0797
SP0798
SP0799
SP0801
SP0802
SP0803
SP0805
SP0806
SP0807
SP0816
SP0817
SP0820
SP0822
SP0823
SP0824
SP0825
SP0828
SP0829
SP0831
SP0835
SP0837
SP0838
SP0839
SP0841
SP0843
SP0844
SP0845
SP0846
SP0847
SP0848
SP0851
SP0852
SP0855
SP0856
SP0862
SP0864
SP0865
SP0867
SP0868
SP0869
SP0870
SP0871
SP0872
SP0873
SP0875
SP0876
SP0877
SP0878
SP0880
SP0881
SP0893
SP0894
SP0895
SP0896
SP0897
SP0904
SP0905
SP0908
SP0909
SP0912
SP0923
SP0927
SP0928
SP0929
SP0931
SP0932
SP0933
SP0935
SP0936
SP0937
SP0938
SP0943
SP0944
SP0945
SP0946
SP0947
SP0948
SP0954
SP0955
SP0959
SP0960
SP0961
SP0962
SP0964
SP0966
SP0967
SP0968
SP0969
SP0970
SP0971
SP0972
SP0974
SP0975
SP0976
SP0978
SP0979
SP0980
SP0981
SP0984
SP0985
SP0987
SP0988
SP0989
SP0991
SP0992
SP0993
SP1002
SP1003

Table 36: Spn ORF's are shared with GBS and GAS

SP1004
SP1008
SP1010
SP1012
SP1016
SP1017
SP1018
SP1020
SP1021
SP1022
SP1024
SP1025
SP1026
SP1029
SP1033
SP1034
SP1035
SP1045
SP1056
SP1067
SP1068
SP1069
SP1070
SP1071
SP1072
SP1073
SP1074
SP1076
SP1079
SP1081
SP1082
SP1083
SP1084
SP1087
SP1088
SP1089
SP1090
SP1093
SP1094
SP1095
SP1096
SP1097
SP1098
SP1099
SP1100
SP1102
SP1105
SP1106
SP1107
SP1110
SP1111
SP1112
SP1113
SP1114
SP1115
SP1116
SP1117
SP1118
SP1119
SP1128
SP1151
SP1152
SP1155
SP1156
SP1157
SP1159
SP1160
SP1161
SP1162
SP1163
SP1164
SP1167
SP1168
SP1169
SP1174
SP1175
SP1176
SP1177
SP1178
SP1179
SP1180
SP1182
SP1184
SP1185
SP1187
SP1190
SP1191
SP1192
SP1193
SP1197
SP1200
SP1202
SP1204
SP1205
SP1207
SP1208
SP1212
SP1213
SP1218
SP1219
SP1220
SP1225
SP1226
SP1227
SP1228
SP1229
SP1230
SP1231
SP1232
SP1233
SP1238
SP1241
SP1242
SP1244
SP1245
SP1246
SP1247
SP1248
SP1249
SP1260
SP1263
SP1266
SP1275
SP1276
SP1277
SP1278
SP1279
SP1280
SP1283
SP1284
SP1285
SP1286
SP1287
SP1288
SP1289
SP1290
SP1291
SP1293
SP1297
SP1298
SP1299
SP1308
SP1316
SP1324
SP1329
SP1330
SP1331
SP1336
SP1341
SP1354
SP1355
SP1357
SP1358
SP1359
SP1362
SP1368
SP1370
SP1371
SP1372
SP1374
SP1375
SP1376
SP1377
SP1378
SP1380
SP1381
SP1383
SP1386
SP1387
SP1388
SP1389
SP1390
SP1393
SP1394
SP1395
SP1396
SP1397
SP1398
SP1399
SP1400
SP1402
SP1403
SP1404
SP1405
SP1406
SP1407
SP1408
SP1409
SP1411
SP1412
SP1413
SP1414
SP1415
SP1416
SP1420
SP1421
SP1427
SP1428
SP1429
SP1434
SP1435
SP1445
SP1446
SP1448
SP1449
SP1450
SP1452
SP1453
SP1456
SP1457
SP1458
SP1460
SP1461
SP1462
SP1465
SP1466
SP1469
SP1470
SP1473
SP1474
SP1475
SP1478
SP1479
SP1482

Table 36: Spn ORF's are shared with GBS and GAS

SP1483
SP1485
SP1489
SP1491
SP1498
SP1500
SP1501
SP1502
SP1504
SP1505
SP1507
SP1508
SP1509
SP1510
SP1511
SP1512
SP1513
SP1517
SP1518
SP1519
SP1521
SP1522
SP1523
SP1529
SP1530
SP1534
SP1535
SP1536
SP1537
SP1538
SP1539
SP1540
SP1541
SP1542
SP1544
SP1547
SP1549
SP1551
SP1552
SP1553
SP1554
SP1557
SP1558
SP1559
SP1560
SP1561
SP1563
SP1564
SP1565
SP1566
SP1568
SP1569
SP1571
SP1574
SP1575
SP1577
SP1580
SP1583
SP1584
SP1586
SP1587
SP1588
SP1589
SP1590
SP1591
SP1597
SP1598
SP1599
SP1602
SP1603
SP1606
SP1608
SP1609
SP1610
SP1615
SP1616
SP1617
SP1624
SP1625
SP1626
SP1631
SP1633
SP1638
SP1644
SP1645
SP1646
SP1647
SP1648
SP1649
SP1650
SP1652
SP1653
SP1655
SP1659
SP1661
SP1662
SP1664
SP1665
SP1666
SP1667
SP1668
SP1670
SP1671
SP1672
SP1674
SP1675
SP1676
SP1677
SP1681
SP1682
SP1683
SP1684
SP1685
SP1688
SP1689
SP1697
SP1698
SP1699
SP1702
SP1709
SP1711
SP1712
SP1713
SP1714
SP1717
SP1721
SP1722
SP1724
SP1725
SP1726
SP1727
SP1732
SP1733
SP1734
SP1735
SP1736
SP1737
SP1738
SP1739
SP1742
SP1743
SP1744
SP1746
SP1747
SP1748
SP1749
SP1750
SP1752
SP1759
SP1776
SP1780
SP1781
SP1782
SP1785
SP1790
SP1795
SP1799
SP1804
SP1816
SP1817
SP1825
SP1839
SP1840
SP1845
SP1847
SP1848
SP1851
SP1855
SP1857
SP1858
SP1860
SP1861
SP1865
SP1871
SP1873
SP1874
SP1875
SP1876
SP1877
SP1878
SP1879
SP1880
SP1881
SP1883
SP1884
SP1887
SP1888
SP1889
SP1890
SP1895
SP1896
SP1900
SP1901
SP1902
SP1903
SP1906
SP1908
SP1909
SP1916
SP1918
SP1922
SP1940
SP1942
SP1944
SP1953
SP1957
SP1960
SP1961
SP1963
SP1964
SP1966
SP1967
SP1968
SP1969
SP1970
SP1972
SP1973
SP1974
SP1975
SP1976
SP1979
SP1980
SP1981
SP1982

Table 36: Spn ORF's are shared with GBS and GAS

SP1983
SP1984
SP1985
SP1987
SP1989
SP1990
SP1991
SP1993
SP1994
SP1996
SP1997
SP1998
SP1999
SP2006
SP2007
SP2010
SP2011
SP2012
SP2020
SP2021
SP2022
SP2027
SP2028
SP2030
SP2031
SP2032
SP2033
SP2034
SP2035
SP2036
SP2037
SP2038
SP2040
SP2041
SP2042
SP2044
SP2045
SP2048
SP2052
SP2053
SP2054
SP2055
SP2056
SP2057
SP2058
SP2063
SP2065
SP2069
SP2070
SP2072
SP2073
SP2075
SP2077
SP2078
SP2082
SP2083
SP2085
SP2086
SP2087
SP2088
SP2090
SP2091
SP2092
SP2094
SP2099
SP2100
SP2101
SP2106
SP2107
SP2108
SP2109
SP2110
SP2112
SP2113
SP2114
SP2119
SP2121
SP2129
SP2131
SP2135
SP2142
SP2148
SP2150
SP2151
SP2152
SP2153
SP2156
SP2161
SP2162
SP2169
SP2170
SP2171
SP2172
SP2173
SP2174
SP2175
SP2176
SP2184
SP2185
SP2186
SP2187
SP2188
SP2189
SP2191
SP2192
SP2193
SP2194
SP2195
SP2202
SP2203
SP2204
SP2205
SP2206
SP2207
SP2208
SP2209
SP2210
SP2214
SP2215
SP2216
SP2219
SP2220
SP2221
SP2222
SP2224
SP2225
SP2226
SP2227
SP2228
SP2229
SP2230
SP2231
SP2233
SP2234
SP2235
SP2238
SP2239
SP2240

Table 37: Spn ORF's which are shared with GBS but not with GAS

SP0012
SP0020
SP0039
SP0050
SP0082
SP0107
SP0113
SP0119
SP0146
SP0150
SP0175
SP0176
SP0177
SP0178
SP0237
SP0255
SP0260
SP0267
SP0278
SP0288
SP0346
SP0347
SP0348
SP0349
SP0366
SP0376
SP0413
SP0445
SP0462
SP0463
SP0479
SP0480
SP0482
SP0484
SP0537
SP0538
SP0566
SP0580
SP0585
SP0599
SP0600
SP0601
SP0606
SP0607
SP0609
SP0617
SP0627
SP0655
SP0656
SP0710
SP0711
SP0717
SP0718
SP0720
SP0723
SP0724
SP0725
SP0730
SP0739
SP0749
SP0750
SP0751
SP0752
SP0753
SP0754
SP0769
SP0789
SP0791
SP0826
SP0900
SP0913
SP0914
SP0939
SP0941
SP0942
SP0953
SP0973
SP0977
SP1011
SP1013
SP1027
SP1054
SP1055
SP1080
SP1086
SP1121
SP1122
SP1123
SP1124
SP1126
SP1127
SP1137
SP1166
SP1173
SP1194
SP1195
SP1215
SP1240
SP1256
SP1261
SP1271
SP1272
SP1273
SP1274
SP1306
SP1310
SP1332
SP1333
SP1334
SP1346
SP1348
SP1350
SP1360
SP1361
SP1365
SP1382
SP1384
SP1392
SP1447
SP1451
SP1463
SP1464
SP1471
SP1472
SP1524
SP1527
SP1600
SP1605
SP1607
SP1632
SP1634
SP1651
SP1673
SP1680
SP1695
SP1700
SP1701
SP1720
SP1729
SP1740
SP1741
SP1745
SP1751
SP1757
SP1758
SP1761
SP1762
SP1763
SP1764
SP1765
SP1766
SP1767
SP1768
SP1770
SP1771
SP1772
SP1783
SP1802
SP1828
SP1856
SP1867
SP1869
SP1870
SP1872
SP1891
SP1907
SP1910
SP1911
SP1927
SP1928
SP1943
SP1959
SP2001
SP2002
SP2009
SP2026
SP2029
SP2039
SP2061
SP2064
SP2066
SP2079
SP2084
SP2095
SP2096
SP2098
SP2103
SP2127
SP2128
SP2130
SP2134
SP2137
SP2138
SP2157
SP2196

Table 38: Spn ORF's which are shared with GAS but no with GBS

SP0065
SP0075
SP0090
SP0091
SP0092
SP0099
SP0100
SP0153
SP0155
SP0156
SP0200
SP0306
SP0313
SP0341
SP0476
SP0496
SP0509
SP0527
SP0648
SP0658
SP0659
SP0661
SP0677
SP0715
SP0742
SP0743
SP0858
SP0859
SP0860
SP0910
SP0986
SP0994
SP0999
SP1000
SP1001
SP1023
SP1075
SP1129
SP1147
SP1171
SP1186
SP1315
SP1317
SP1319
SP1320
SP1321
SP1322
SP1438
SP1442
SP1525
SP1546
SP1570
SP1572
SP1578
SP1604
SP1715
SP1754
SP1797
SP1798
SP1800
SP1885
SP1919
SP1923
SP1941
SP1950
SP2016
SP2017
SP2051
SP2060
SP2111
SP2143
SP2144
SP2201
SP2236

Table 40: Comparative Sequences relating to SAG0635

SEQ ID NO 4001 : SAG0653 FROM THE 2603 V/R GBS TYPE V STRAIN
ATGAAGAAAGTGTTAGTGAGTAGTCTTTTGGTTTTAGGGATTACGATA
ACGTTACAAACAGTAGTTGAGGCTAAGGGGCCAAAAGTAGCTTATACACAAGAGGGAATG
ACTGCTCTTTCGGACACAAATAAAGATAAAGTCACTACTATTTCTATTGACGAGATTCAA
AAAAGCTTAGAAGGTAAGAAGCCGATTACTGTTAGTTTTGATATTGATGATACACTGCTT
TTCAGTAGTCAATATTTTCAATATGGTAAAGAATATGTAACTCCTGGATCGTTTGATTTT
CTTCATAAACAAAAATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCATTCCC
AAAGAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAGATAAAATTGTTTTT
ATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAGGGCGAGGTTGATAAAACAGCTAAA
GCCTTAGCTAAAGATTTTAAATTAGACAAACCAATTGCTGTAAATTATACAGGCGATAAA
CCTAAAAAGCCATACAAATATGATAAATCATATTATATTAAGAAATATGGTTCAGACATT
CATTATGGAGATAGTGATGACGATATTCATGCAGCTAGGGAGGCCGGTGCTAGACCAATT
AGAATTTTAAGAGCACCTAATTCTACAAATCTACCTTTACCAGAAGCTGGAGGCTACGGT
GAAGAGGTTCTCGAAAATTCAGCTTAC

SEQ ID NO 4002 : SAG0653 FROM THE 090 GBS TYPE III STRAIN
AAGGGGCCAAAAGTAGCTTATACACAAGAGGGAATGAC
TGCTCTTTCGGACACAAATAAAGATAAAGTCACTACTATTTCTATTGACG
AGATTCAAAAAAGCTTAGAAGGTAAGAAGCCGATTACTGTTAGTTTTGAT
ATTGATGATACACTACTTTTCAGTAGTCAATATTTTCAATATGGTAAAGA
ATATGTAACTCCTGGATCGTTTGATTTTCTTCATAAACAAAAATTCTGGG
ATCTTGTTGCAAAACGAGGAGATCAAGATTCCATTCCCAAAGAATATGCT
AAAAAATTAATTGCTATGCATCAAAAACGAGGAGATAAAATTGTTTTTAT
AACAGGTAGGACAAGAGGGTCAATGTATAAGGAGGGCGAGGTTGATAAAA
CAGCTAAAGCCTTAGCTAAAGATTTTAAATTAGACAAACCAATTGCTGTA
AATTATACAGGCGATAAACCTAAAAAGCCATACAAATATGATAAATCATA
TTATATTAAGAAATATGGTTCAGACATTCATTATGGAGATAGTGATGACG
ATATTCATGCAGCTAGGGAGGCCGGTGCTAGACCAATTAGAATTTTAAGA
GCACCTAATTCTACAAATCTACCTTTACCAGAAGCTGGAGGCTACGGTGA
AGAGGTTCTCGAAAATTCAGCTTAC

SEQ ID NO 4003 : SAG0653 FROM THE A909 GBS TYPE Ia STRAIN
AAGGGGCCAAAAGTAGCTTATACACA
AGAGGGAATGACTGCTCTTTCGGACACAAATAAAGATAAAGTCACTACTA
TTTCTATTGACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGCCGATTACT
GTTAGTTTTGATATTGATGATACACTGCTTTTCAGTAGTCAATATTTTCA
ATATGGTAAAGAATATGTAACTCCTGGATCGTTTGATTTTCTTCATAAAC
AAAAATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCATTCCC
AAAGAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAGATAA
AATTGTTTTTATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAGGGCG
AGGTTGATAAAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTAGACAAA
CCAATTGCTGTAAATTATACAGGCGATAAACCTAAAAAGCCATACAAATA
TGATAAATCATATTATATTAAGAAATATGGTTCAGACATTCATTATGGAG
ATAGTGATGACGATATTCATGCAGCTAGGGAGGCCGGTGCTAGACCAATT
AGAATTTTAAGAGCACCTAATTCTACAAATCTACCTTTACCAGAAGCTGG
AGGCTACGGTGAAGAGGTTCTCGAAAATTCAGCTTAC

SEQ ID NO 4004 : SAG0653 FROM THE 18RS21 GBS TYPE II STRAIN
AAGGGGCCAAAAGTAGCTTATACACAAGA
GGGAATGACTGCTCTTTCGGACACAAATAAAGATAAAGTCACTACTATTT
CTATTGACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGCCGATTACTGTT
AGTTTTGATATTGATGATACACTGCTTTTCAGTAGTCAATATTTTCAATA
TGGTAAAGAATATGTAACTCCTGGATCGTTTGATTTTCTTCATAAACAAA
AATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCATTCCCAAA
GAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAGATAAAAT
TGTTTTTATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAGGGCGAGG
TTGATAAAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTAGACAAACCA
ATTGCTGTAAATTATACAGGCGATAAACCTAAAAAGCCATACAAATATGA
TAAATCATATTATATTAAGAAATATGGTTCAGACATTCATTATGGAGATA
GTGATGACGATATTCATGCAGCTAGGGAGGCCGGTGCTAGACCAATTAGA
ATTTTAAGAGCACCTAATTCTACAAATCTACCTTTACCAGAAGCTGGAGG
CTACGGTGAAGAGGTTCTCGAAAATTCAGCTTAC

SEQ ID NO 4005 : SAG0653 FROM THE M732 GBS TYPE III STRAIN
AAGGGGCCAAAAGTAGCTTATACACAAGA
GGGAATGACTGCTCTTTCGGACACAAATAAAGATAAAGTCACTACTATTT
CTATTGACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGCCGATTACTGTT
AGTTTTGATATTGATGATACACTGCTTTTCAGTAGTCAATATTTTCAATA
TGGTAAAGAATATGTAACTCCTGGATCGTTTGATTTTCTTCATAAACAAA
AATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCATTCCCAAA
GAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAGATAAAAT
TGTTTTTATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAGGGCGAGG
TTGATAAAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTAGACAAACCA
ATTGCTGTAAATTATACAGGCGATAAACCTAAAAAGCCATACAAATATGA
TAAATCATATTATATTAAGAAATATGGTTCAGACATTCATTATGGAGATA
GTGATGACGATATTCATGCAGCTAGGGAGGCCGGTGCTAGACCAATTAGA
ATTTTAAGAGCACCTAATTCTACAAATCTACCTTTACCAGAAGCTGGAGG
CTACGGTGAAGAGGTTCTCGAAAATTCAGCTTAC

SEQ ID NO 4006 : SAG0653 FROM THE COH1 GBS TYPE III STRAIN
AAGGGGCCAAAAGTAGCTTATACACAAGAGGGAATGACT
GCTCTTTCGGACACAAATAAAGATAAAGTCACTACTATTTCTATTGACGA
GATTCAAAAAAGCTTAGAAGGTAAGAAGCCGATTACTGTTAGTTTTGATA

Table 40: Comparative Sequences relating to SAG0635

TTGATGATACACTGCTTTTCAGTAGTCAATATTTTCAATATGGTAAAGAA
TATGTAACTCCTGGATCGTTTGATTTTCTTCATAAACAAAAATTCTGGGA
TCTTGTTGCAAAACGAGGAGATCAAGATTCCATTCCCAAAGAATATGCTA
AAAAATTAATTGCTATGCATCAAAAACGAGGAGATAAAATTGTTTTTATA
ACAGGTAGGACAAGAGGGTCAATGTATAAGGAGGGCGAGGTTGATAAAAC
AGCTAAAGCCTTAGCTAAAGATTTTAAATTAGACAAACCAATTGCTGTAA
ATTATACAGGCGATAAACCTAAAAAGCCATACAAATATGATAAATCATAT
TATATTAAGAAATATGGTTCAGACATTCATTATGGAGATAGTGATGACGA
TATTCATGCAGCTAGGGAGGCCGGTGCTAGACCAATTAGAATTTTAAGAG
CACCTAATTCTACAAATCTACCTTTACCAGAAGCTGGAGGCTACGGTGAA
GAGGTTCTCGAAAATTCAGCTTAC

SEQ ID NO 4007 : SAG0653 FROM THE M781 GBS TYPE III STRAIN
AAGGGGCCAAAAGTAGCTTATACACA
AGAGGGAATGACTGCTCTTTCGGACACAAATAAAGATAAAGTCACTACTA
TTTCTATTGACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGCCGATTACT
GTTAGTTTTGATATTGATGATACACTGCTTTTCAGTAGTCAATATTTTCA
ATATGGTAAAGAATATGTAACTCCTGGATCGTTTGATTTTCTTCATAAAC
AAAAATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCATTCCC
AAAGAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAGATAA
AATTGTTTTTATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAGGGCG
AGGTTGATAAAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTAGACAAA
CCAATTGCTGTAAATTATACAGGCGATAAACCTAAAAAGCCATACAAATA
TGATAAATCATATTATATTAAGAAATATGGTTCAGACATTCATTATGGAG
ATAGTGATGACGATATTCATGCAGCTAGGGAGGCCGGTGCTAGACCAATT
AGAATTTTAAGAGCACCTAATTCTACAAATCTACCTTTACCAGAAGCTGG
AGGCTACGGTGAAGAGGTTCTCGAAAATTCAGCTTAC

SEQ ID NO 4008 : SAG0653 FROM THE CJB110 GBS NONTYPEABLE STRAIN
AAGGGGCCAAAAGTAGCTTATACACAAGA
GGGAATGACTGCTCTTTCGGACACAAATAAAGATAAAGTCACTACTATTT
CTATTGACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGCCGATTACTGTT
AGTTTTGATATTGATGATACACTGCTTTTCAGTAGTCAATATTTTCAATA
TGGTAAAGAATATGTAACTCCTGGATCGTTTGATTTTCTTCATAAACAAA
AATTCTGGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCATTCCCAAA
GAATATGCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAGATAAAAT
TGTTTTTATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAGGGCGAGG
TTGATAAAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTAGACAAACCA
ATTGCTGTAAATTATACAGGCGATAAACCTAAAAAGCCATACAAATATGA
TAAATCATATTATATTAAGAAATATGGTTCAGACATTCATTATGGAGATA
GTGATGACGATATTCATGCAGCTAGGGAGGCCGGTGCTAGACCAATTAGA
ATTTTAAGAGCACCTAATTCTACAAATCTACCTTTACCAGAAGCTGGAGG
CTACGGTGAAGAGGTTCTCGAAAATTCAGCTTAC

SEQ ID NO 4009 : SAG0653 FROM THE JM9130013 GBS TYPE VIII STRAIN
AAGGGGCCAAAAGTAGCTTATACACAAGAGGGAAT
GACTGCTCTTTCGGACACAAATAAAGATAAAGTCACTACTATTTCTATTG
ACGAGATTCAAAAAAGCTTAGAAGGTAAGAAGCCGATTACTGTTAGTTTT
GATATTGATGATACACTGCTTTTCAGTAGTCAATATTTTCAATATGGTAA
AGAATATGTAACTCCTGGATCGTTTGATTTTCTTCATAAACAAAAATTCT
GGGATCTTGTTGCAAAACGAGGAGATCAAGATTCCATTCCCAAAGAATAT
GCTAAAAAATTAATTGCTATGCATCAAAAACGAGGAGATAAAATTGTTTT
TATAACAGGTAGGACAAGAGGGTCAATGTATAAGGAGGGCGAGGTTGATA
AAACAGCTAAAGCCTTAGCTAAAGATTTTAAATTAGACAAACCAATTGCT
GTAAATTATACAGGCGATAAACCTAAAAAGCCATACAAATATGATAAATC
ATATTATATTAAGAAATATGGTTCAGACATTCATTATGGAGATAGTGATG
ACGATATTCATGCAGCTAGGGAGGCCGGTGCTAGACCAATTAGAATTTTA
AGAGCACCTAATTCTACAAATCTACCTTTACCAGAAGCTGGAGGCTACGG
TGAAGAGGTTCTCGAAAATTCAGCTTAC

Table 40: Comparative Sequences relating to SAG0635

```
PRETTY of: /biotmp/msa20031.2{*}   August 5, 2002 07:05  ..

                              1                                                  50
    msa20031.2{100_18RS21}    AAGGGGCCAA AAGTAGCTTA TACACAAGAG GGAATGACTG CTCTTTCGGA
      msa20031.2{100_2603}    AAGGGGCCAA AAGTAGCTTA TACACAAGAG GGAATGACTG CTCTTTCGGA
      msa20031.2{100_A909}    AAGGGGCCAA AAGTAGCTTA TACACAAGAG GGAATGACTG CTCTTTCGGA
    msa20031.2{100_CJB110}    AAGGGGCCAA AAGTAGCTTA TACACAAGAG GGAATGACTG CTCTTTCGGA
      msa20031.2{100_COH1}    AAGGGGCCAA AAGTAGCTTA TACACAAGAG GGAATGACTG CTCTTTCGGA
 msa20031.2{100_JM9130013}    AAGGGGCCAA AAGTAGCTTA TACACAAGAG GGAATGACTG CTCTTTCGGA
      msa20031.2{100_M732}    AAGGGGCCAA AAGTAGCTTA TACACAAGAG GGAATGACTG CTCTTTCGGA
      msa20031.2{100_M781}    AAGGGGCCAA AAGTAGCTTA TACACAAGAG GGAATGACTG CTCTTTCGGA
       msa20031.2{100_090}    AAGGGGCCAA AAGTAGCTTA TACACAAGAG GGAATGACTG CTCTTTCGGA
                 Consensus    ********** ********** ********** ********** **********

                              51                                                100
    msa20031.2{100_18RS21}    CACAAATAAA GATAAAGTCA CTACTATTTC TATTGACGAG ATTCAAAAAA
      msa20031.2{100_2603}    CACAAATAAA GATAAAGTCA CTACTATTTC TATTGACGAG ATTCAAAAAA
      msa20031.2{100_A909}    CACAAATAAA GATAAAGTCA CTACTATTTC TATTGACGAG ATTCAAAAAA
    msa20031.2{100_CJB110}    CACAAATAAA GATAAAGTCA CTACTATTTC TATTGACGAG ATTCAAAAAA
      msa20031.2{100_COH1}    CACAAATAAA GATAAAGTCA CTACTATTTC TATTGACGAG ATTCAAAAAA
 msa20031.2{100_JM9130013}    CACAAATAAA GATAAAGTCA CTACTATTTC TATTGACGAG ATTCAAAAAA
      msa20031.2{100_M732}    CACAAATAAA GATAAAGTCA CTACTATTTC TATTGACGAG ATTCAAAAAA
      msa20031.2{100_M781}    CACAAATAAA GATAAAGTCA CTACTATTTC TATTGACGAG ATTCAAAAAA
       msa20031.2{100_090}    CACAAATAAA GATAAAGTCA CTACTATTTC TATTGACGAG ATTCAAAAAA
                 Consensus    ********** ********** ********** ********** **********

                              101                                               150
    msa20031.2{100_18RS21}    GCTTAGAAGG TAAGAAGCCG ATTACTGTTA GTTTTGATAT TGATGATACA
      msa20031.2{100_2603}    GCTTAGAAGG TAAGAAGCCG ATTACTGTTA GTTTTGATAT TGATGATACA
      msa20031.2{100_A909}    GCTTAGAAGG TAAGAAGCCG ATTACTGTTA GTTTTGATAT TGATGATACA
    msa20031.2{100_CJB110}    GCTTAGAAGG TAAGAAGCCG ATTACTGTTA GTTTTGATAT TGATGATACA
      msa20031.2{100_COH1}    GCTTAGAAGG TAAGAAGCCG ATTACTGTTA GTTTTGATAT TGATGATACA
 msa20031.2{100_JM9130013}    GCTTAGAAGG TAAGAAGCCG ATTACTGTTA GTTTTGATAT TGATGATACA
      msa20031.2{100_M732}    GCTTAGAAGG TAAGAAGCCG ATTACTGTTA GTTTTGATAT TGATGATACA
      msa20031.2{100_M781}    GCTTAGAAGG TAAGAAGCCG ATTACTGTTA GTTTTGATAT TGATGATACA
       msa20031.2{100_090}    GCTTAGAAGG TAAGAAGCCG ATTACTGTTA GTTTTGATAT TGATGATACA
                 Consensus    ********** ********** ********** ********** **********

                              151                                               200
    msa20031.2{100_18RS21}    CTgCTTTTCA GTAGTCAATA TTTTCAATAT GGTAAAGAAT ATGTAACTCC
      msa20031.2{100_2603}    CTgCTTTTCA GTAGTCAATA TTTTCAATAT GGTAAAGAAT ATGTAACTCC
      msa20031.2{100_A909}    CTgCTTTTCA GTAGTCAATA TTTTCAATAT GGTAAAGAAT ATGTAACTCC
    msa20031.2{100_CJB110}    CTgCTTTTCA GTAGTCAATA TTTTCAATAT GGTAAAGAAT ATGTAACTCC
      msa20031.2{100_COH1}    CTgCTTTTCA GTAGTCAATA TTTTCAATAT GGTAAAGAAT ATGTAACTCC
 msa20031.2{100_JM9130013}    CTgCTTTTCA GTAGTCAATA TTTTCAATAT GGTAAAGAAT ATGTAACTCC
      msa20031.2{100_M732}    CTgCTTTTCA GTAGTCAATA TTTTCAATAT GGTAAAGAAT ATGTAACTCC
      msa20031.2{100_M781}    CTgCTTTTCA GTAGTCAATA TTTTCAATAT GGTAAAGAAT ATGTAACTCC
       msa20031.2{100_090}    CTaCTTTTCA GTAGTCAATA TTTTCAATAT GGTAAAGAAT ATGTAACTCC
                 Consensus    **-******* ********** ********** ********** **********

                              201                                               250
    msa20031.2{100_18RS21}    TGGATCGTTT GATTTTCTTC ATAAACAAAA ATTCTGGGAT CTTGTTGCAA
      msa20031.2{100_2603}    TGGATCGTTT GATTTTCTTC ATAAACAAAA ATTCTGGGAT CTTGTTGCAA
      msa20031.2{100_A909}    TGGATCGTTT GATTTTCTTC ATAAACAAAA ATTCTGGGAT CTTGTTGCAA
    msa20031.2{100_CJB110}    TGGATCGTTT GATTTTCTTC ATAAACAAAA ATTCTGGGAT CTTGTTGCAA
      msa20031.2{100_COH1}    TGGATCGTTT GATTTTCTTC ATAAACAAAA ATTCTGGGAT CTTGTTGCAA
 msa20031.2{100_JM9130013}    TGGATCGTTT GATTTTCTTC ATAAACAAAA ATTCTGGGAT CTTGTTGCAA
      msa20031.2{100_M732}    TGGATCGTTT GATTTTCTTC ATAAACAAAA ATTCTGGGAT CTTGTTGCAA
      msa20031.2{100_M781}    TGGATCGTTT GATTTTCTTC ATAAACAAAA ATTCTGGGAT CTTGTTGCAA
       msa20031.2{100_090}    TGGATCGTTT GATTTTCTTC ATAAACAAAA ATTCTGGGAT CTTGTTGCAA
                 Consensus    ********** ********** ********** ********** **********

                              251                                               300
    msa20031.2{100_18RS21}    AACGAGGAGA TCAAGATTCC ATTCCCAAAG AATATGCTAA AAAATTAATT
      msa20031.2{100_2603}    AACGAGGAGA TCAAGATTCC ATTCCCAAAG AATATGCTAA AAAATTAATT
      msa20031.2{100_A909}    AACGAGGAGA TCAAGATTCC ATTCCCAAAG AATATGCTAA AAAATTAATT
    msa20031.2{100_CJB110}    AACGAGGAGA TCAAGATTCC ATTCCCAAAG AATATGCTAA AAAATTAATT
      msa20031.2{100_COH1}    AACGAGGAGA TCAAGATTCC ATTCCCAAAG AATATGCTAA AAAATTAATT
 msa20031.2{100_JM9130013}    AACGAGGAGA TCAAGATTCC ATTCCCAAAG AATATGCTAA AAAATTAATT
      msa20031.2{100_M732}    AACGAGGAGA TCAAGATTCC ATTCCCAAAG AATATGCTAA AAAATTAATT
      msa20031.2{100_M781}    AACGAGGAGA TCAAGATTCC ATTCCCAAAG AATATGCTAA AAAATTAATT
       msa20031.2{100_090}    AACGAGGAGA TCAAGATTCC ATTCCCAAAG AATATGCTAA AAAATTAATT
                 Consensus    ********** ********** ********** ********** **********

                              301                                               350
    msa20031.2{100_18RS21}    GCTATGCATC AAAAACGAGG AGATAAAATT GTTTTTATAA CAGGTAGGAC
      msa20031.2{100_2603}    GCTATGCATC AAAAACGAGG AGATAAAATT GTTTTTATAA CAGGTAGGAC
      msa20031.2{100_A909}    GCTATGCATC AAAAACGAGG AGATAAAATT GTTTTTATAA CAGGTAGGAC
    msa20031.2{100_CJB110}    GCTATGCATC AAAAACGAGG AGATAAAATT GTTTTTATAA CAGGTAGGAC
      msa20031.2{100_COH1}    GCTATGCATC AAAAACGAGG AGATAAAATT GTTTTTATAA CAGGTAGGAC
 msa20031.2{100_JM9130013}    GCTATGCATC AAAAACGAGG AGATAAAATT GTTTTTATAA CAGGTAGGAC
      msa20031.2{100_M732}    GCTATGCATC AAAAACGAGG AGATAAAATT GTTTTTATAA CAGGTAGGAC
      msa20031.2{100_M781}    GCTATGCATC AAAAACGAGG AGATAAAATT GTTTTTATAA CAGGTAGGAC
```

Table 40: Comparative Sequences relating to SAG0635

```
       msa20031.2{100_090}  GCTATGCATC AAAAACGAGG AGATAAAATT GTTTTTATAA CAGGTAGGAC
                 Consensus  ********** ********** ********** ********** **********

                            351                                               400
     msa20031.2{100_18RS21}  AAGAGGGTCA ATGTATAAGG AGGGCGAGGT TGATAAAACA GCTAAAGCCT
       msa20031.2{100_2603}  AAGAGGGTCA ATGTATAAGG AGGGCGAGGT TGATAAAACA GCTAAAGCCT
       msa20031.2{100_A909}  AAGAGGGTCA ATGTATAAGG AGGGCGAGGT TGATAAAACA GCTAAAGCCT
     msa20031.2{100_CJB110}  AAGAGGGTCA ATGTATAAGG AGGGCGAGGT TGATAAAACA GCTAAAGCCT
       msa20031.2{100_COH1}  AAGAGGGTCA ATGTATAAGG AGGGCGAGGT TGATAAAACA GCTAAAGCCT
  msa20031.2{100_JM9130013}  AAGAGGGTCA ATGTATAAGG AGGGCGAGGT TGATAAAACA GCTAAAGCCT
       msa20031.2{100_M732}  AAGAGGGTCA ATGTATAAGG AGGGCGAGGT TGATAAAACA GCTAAAGCCT
       msa20031.2{100_M781}  AAGAGGGTCA ATGTATAAGG AGGGCGAGGT TGATAAAACA GCTAAAGCCT
        msa20031.2{100_090}  AAGAGGGTCA ATGTATAAGG AGGGCGAGGT TGATAAAACA GCTAAAGCCT
                  Consensus  ********** ********** ********** ********** **********

                            401                                               450
     msa20031.2{100_18RS21}  TAGCTAAAGA TTTTAAATTA GACAAACCAA TTGCTGTAAA TTATACAGGC
       msa20031.2{100_2603}  TAGCTAAAGA TTTTAAATTA GACAAACCAA TTGCTGTAAA TTATACAGGC
       msa20031.2{100_A909}  TAGCTAAAGA TTTTAAATTA GACAAACCAA TTGCTGTAAA TTATACAGGC
     msa20031.2{100_CJB110}  TAGCTAAAGA TTTTAAATTA GACAAACCAA TTGCTGTAAA TTATACAGGC
       msa20031.2{100_COH1}  TAGCTAAAGA TTTTAAATTA GACAAACCAA TTGCTGTAAA TTATACAGGC
  msa20031.2{100_JM9130013}  TAGCTAAAGA TTTTAAATTA GACAAACCAA TTGCTGTAAA TTATACAGGC
       msa20031.2{100_M732}  TAGCTAAAGA TTTTAAATTA GACAAACCAA TTGCTGTAAA TTATACAGGC
       msa20031.2{100_M781}  TAGCTAAAGA TTTTAAATTA GACAAACCAA TTGCTGTAAA TTATACAGGC
        msa20031.2{100_090}  TAGCTAAAGA TTTTAAATTA GACAAACCAA TTGCTGTAAA TTATACAGGC
                  Consensus  ********** ********** ********** ********** **********

                            451                                               500
     msa20031.2{100_18RS21}  GATAAACCTA AAAAGCCATA CAAATATGAT AAATCATATT ATATTAAGAA
       msa20031.2{100_2603}  GATAAACCTA AAAAGCCATA CAAATATGAT AAATCATATT ATATTAAGAA
       msa20031.2{100_A909}  GATAAACCTA AAAAGCCATA CAAATATGAT AAATCATATT ATATTAAGAA
     msa20031.2{100_CJB110}  GATAAACCTA AAAAGCCATA CAAATATGAT AAATCATATT ATATTAAGAA
       msa20031.2{100_COH1}  GATAAACCTA AAAAGCCATA CAAATATGAT AAATCATATT ATATTAAGAA
  msa20031.2{100_JM9130013}  GATAAACCTA AAAAGCCATA CAAATATGAT AAATCATATT ATATTAAGAA
       msa20031.2{100_M732}  GATAAACCTA AAAAGCCATA CAAATATGAT AAATCATATT ATATTAAGAA
       msa20031.2{100_M781}  GATAAACCTA AAAAGCCATA CAAATATGAT AAATCATATT ATATTAAGAA
        msa20031.2{100_090}  GATAAACCTA AAAAGCCATA CAAATATGAT AAATCATATT ATATTAAGAA
                  Consensus  ********** ********** ********** ********** **********

                            501                                               550
     msa20031.2{100_18RS21}  ATATGGTTCA GACATTCATT ATGGAGATAG TGATGACGAT ATTCATGCAG
       msa20031.2{100_2603}  ATATGGTTCA GACATTCATT ATGGAGATAG TGATGACGAT ATTCATGCAG
       msa20031.2{100_A909}  ATATGGTTCA GACATTCATT ATGGAGATAG TGATGACGAT ATTCATGCAG
     msa20031.2{100_CJB110}  ATATGGTTCA GACATTCATT ATGGAGATAG TGATGACGAT ATTCATGCAG
       msa20031.2{100_COH1}  ATATGGTTCA GACATTCATT ATGGAGATAG TGATGACGAT ATTCATGCAG
  msa20031.2{100_JM9130013}  ATATGGTTCA GACATTCATT ATGGAGATAG TGATGACGAT ATTCATGCAG
       msa20031.2{100_M732}  ATATGGTTCA GACATTCATT ATGGAGATAG TGATGACGAT ATTCATGCAG
       msa20031.2{100_M781}  ATATGGTTCA GACATTCATT ATGGAGATAG TGATGACGAT ATTCATGCAG
        msa20031.2{100_090}  ATATGGTTCA GACATTCATT ATGGAGATAG TGATGACGAT ATTCATGCAG
                  Consensus  ********** ********** ********** ********** **********

                            551                                               600
     msa20031.2{100_18RS21}  CTAGGGAGGC CGGTGCTAGA CCAATTAGAA TTTTAAGAGC ACCTAATTCT
       msa20031.2{100_2603}  CTAGGGAGGC CGGTGCTAGA CCAATTAGAA TTTTAAGAGC ACCTAATTCT
       msa20031.2{100_A909}  CTAGGGAGGC CGGTGCTAGA CCAATTAGAA TTTTAAGAGC ACCTAATTCT
     msa20031.2{100_CJB110}  CTAGGGAGGC CGGTGCTAGA CCAATTAGAA TTTTAAGAGC ACCTAATTCT
       msa20031.2{100_COH1}  CTAGGGAGGC CGGTGCTAGA CCAATTAGAA TTTTAAGAGC ACCTAATTCT
  msa20031.2{100_JM9130013}  CTAGGGAGGC CGGTGCTAGA CCAATTAGAA TTTTAAGAGC ACCTAATTCT
       msa20031.2{100_M732}  CTAGGGAGGC CGGTGCTAGA CCAATTAGAA TTTTAAGAGC ACCTAATTCT
       msa20031.2{100_M781}  CTAGGGAGGC CGGTGCTAGA CCAATTAGAA TTTTAAGAGC ACCTAATTCT
        msa20031.2{100_090}  CTAGGGAGGC CGGTGCTAGA CCAATTAGAA TTTTAAGAGC ACCTAATTCT
                  Consensus  ********** ********** ********** ********** **********

                            601                                               650
     msa20031.2{100_18RS21}  ACAAATCTAC CTTTACCAGA AGCTGGAGGC TACGGTGAAG AGGTTCTCGA
       msa20031.2{100_2603}  ACAAATCTAC CTTTACCAGA AGCTGGAGGC TACGGTGAAG AGGTTCTCGA
       msa20031.2{100_A909}  ACAAATCTAC CTTTACCAGA AGCTGGAGGC TACGGTGAAG AGGTTCTCGA
     msa20031.2{100_CJB110}  ACAAATCTAC CTTTACCAGA AGCTGGAGGC TACGGTGAAG AGGTTCTCGA
       msa20031.2{100_COH1}  ACAAATCTAC CTTTACCAGA AGCTGGAGGC TACGGTGAAG AGGTTCTCGA
  msa20031.2{100_JM9130013}  ACAAATCTAC CTTTACCAGA AGCTGGAGGC TACGGTGAAG AGGTTCTCGA
       msa20031.2{100_M732}  ACAAATCTAC CTTTACCAGA AGCTGGAGGC TACGGTGAAG AGGTTCTCGA
       msa20031.2{100_M781}  ACAAATCTAC CTTTACCAGA AGCTGGAGGC TACGGTGAAG AGGTTCTCGA
        msa20031.2{100_090}  ACAAATCTAC CTTTACCAGA AGCTGGAGGC TACGGTGAAG AGGTTCTCGA
                  Consensus  ********** ********** ********** ********** **********

                            651        663
     msa20031.2{100_18RS21}  AAATTCAGCT TAC
       msa20031.2{100_2603}  AAATTCAGCT TAC
       msa20031.2{100_A909}  AAATTCAGCT TAC
     msa20031.2{100_CJB110}  AAATTCAGCT TAC
       msa20031.2{100_COH1}  AAATTCAGCT TAC
  msa20031.2{100_JM9130013}  AAATTCAGCT TAC
       msa20031.2{100_M732}  AAATTCAGCT TAC
```

Table 40: Comparative Sequences relating to SAG0635

```
msa20031.2{100_M781}  AAATTCAGCT TAC
 msa20031.2{100_090}  AAATTCAGCT TAC
           Consensus  ********** ***
```

SEQ ID NO 4010 : SAG0653 FROM THE 2603 V/R GBS TYPE V STRAIN
KGPKVAYTQEGMTALSDTNKDKVTTISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQY
GKEYVTPGSFDFLHKQKFWDLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGS
MYKEGEVDKTAKALAKDFKLDKPIAVNYTGDKPKKPYKYDKSYYIKKYGSDIHYGDSDDD
IHAAREAGARPIRILRAPNSTNLPLPEAGGYGEEVLENSAY

SEQ ID NO 4011 : SAG0653 FROM THE 090 GBS TYPE III STRAIN
KGPKVAYTQEGMTALSDTNKDKVTTISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQY
GKEYVTPGSFDFLHKQKFWDLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGS
MYKEGEVDKTAKALAKDFKLDKPIAVNYTGDKPKKPYKYDKSYYIKKYGSDIHYGDSDDD
IHAAREAGARPIRILRAPNSTNLPLPEAGGYGEEVLENSAY

SEQ ID NO 4012 : SAG0653 FROM THE A909 GBS TYPE Ia STRAIN
KGPKVAYTQEGMTALSDTNKDKVTTISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQY
GKEYVTPGSFDFLHKQKFWDLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGS
MYKEGEVDKTAKALAKDFKLDKPIAVNYTGDKPKKPYKYDKSYYIKKYGSDIHYGDSDDD
IHAAREAGARPIRILRAPNSTNLPLPEAGGYGEEVLENSAY

SEQ ID NO 4013 : SAG0653 FROM THE 18RS21 GBS TYPE II STRAIN
KGPKVAYTQEGMTALSDTNKDKVTTISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQY
GKEYVTPGSFDFLHKQKFWDLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGS
MYKEGEVDKTAKALAKDFKLDKPIAVNYTGDKPKKPYKYDKSYYIKKYGSDIHYGDSDDD
IHAAREAGARPIRILRAPNSTNLPLPEAGGYGEEVLENSAY

SEQ ID NO 4014 : SAG0653 FROM THE COH1 GBS TYPE III STRAIN
KGPKVAYTQEGMTALSDTNKDKVTTISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQY
GKEYVTPGSFDFLHKQKFWDLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGS
MYKEGEVDKTAKALAKDFKLDKPIAVNYTGDKPKKPYKYDKSYYIKKYGSDIHYGDSDDD
IHAAREAGARPIRILRAPNSTNLPLPEAGGYGEEVLENSAY

SEQ ID NO 4015 : SAG0653 FROM THE M781 GBS TYPE III STRAIN
KGPKVAYTQEGMTALSDTNKDKVTTISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQY
GKEYVTPGSFDFLHKQKFWDLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGS
MYKEGEVDKTAKALAKDFKLDKPIAVNYTGDKPKKPYKYDKSYYIKKYGSDIHYGDSDDD
IHAAREAGARPIRILRAPNSTNLPLPEAGGYGEEVLENSAY

SEQ ID NO 4016 : SAG0653 FROM THE CJB110 GBS NONTYPEABLE STRAIN
KGPKVAYTQEGMTALSDTNKDKVTTISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQY
GKEYVTPGSFDFLHKQKFWDLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGS
MYKEGEVDKTAKALAKDFKLDKPIAVNYTGDKPKKPYKYDKSYYIKKYGSDIHYGDSDDD
IHAAREAGARPIRILRAPNSTNLPLPEAGGYGEEVLENSAY

SEQ ID NO 4017 : SAG0653 FROM THE JM9130013 GBS TYPE VIII STRAIN
KGPKVAYTQEGMTALSDTNKDKVTTISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQY
GKEYVTPGSFDFLHKQKFWDLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGS
MYKEGEVDKTAKALAKDFKLDKPIAVNYTGDKPKKPYKYDKSYYIKKYGSDIHYGDSDDD
IHAAREAGARPIRILRAPNSTNLPLPEAGGYGEEVLENSAY

SEQ ID NO 4018 : SAG0653 FROM THE M732 GBS TYPE III STRAIN
KGPKVAYTQEGMTALSDTNKDKVTTISIDEIQKSLEGKKPITVSFDIDDTLLFSSQYFQY
GKEYVTPGSFDFLHKQKFWDLVAKRGDQDSIPKEYAKKLIAMHQKRGDKIVFITGRTRGS
MYKEGEVDKTAKALAKDFKLDKPIAVNYTGDKPKKPYKYDKSYYIKKYGSDIHYGDSDDD
IHAAREAGARPIRILRAPNSTNLPLPEAGGYGEEVLENSAY

Table 40: Comparative Sequences relating to SAG0635

```
PRETTY of: /biotmp/msa25122.2{*}   August 5, 2002 07:09  ..

                               1                                                  50
        msa25122.2{100_090}  KGPKVAYTQE GMTALSDTNK DKVTTISIDE IQKSLEGKKP ITVSFDIDDT
      msa25122.2{100_18RS21}  KGPKVAYTQE GMTALSDTNK DKVTTISIDE IQKSLEGKKP ITVSFDIDDT
        msa25122.2{100_2603}  KGPKVAYTQE GMTALSDTNK DKVTTISIDE IQKSLEGKKP ITVSFDIDDT
        msa25122.2{100_A909}  KGPKVAYTQE GMTALSDTNK DKVTTISIDE IQKSLEGKKP ITVSFDIDDT
      msa25122.2{100_CJB110}  KGPKVAYTQE GMTALSDTNK DKVTTISIDE IQKSLEGKKP ITVSFDIDDT
        msa25122.2{100_COH1}  KGPKVAYTQE GMTALSDTNK DKVTTISIDE IQKSLEGKKP ITVSFDIDDT
   msa25122.2{100_JM9130013}  KGPKVAYTQE GMTALSDTNK DKVTTISIDE IQKSLEGKKP ITVSFDIDDT
        msa25122.2{100_M732}  KGPKVAYTQE GMTALSDTNK DKVTTISIDE IQKSLEGKKP ITVSFDIDDT
        msa25122.2{100_M781}  KGPKVAYTQE GMTALSDTNK DKVTTISIDE IQKSLEGKKP ITVSFDIDDT
                  Consensus  ********** ********** ********** ********** **********

                             51                                                  100
        msa25122.2{100_090}  LLFSSQYFQY GKEYVTPGSF DFLHKQKFWD LVAKRGDQDS IPKEYAKKLI
      msa25122.2{100_18RS21}  LLFSSQYFQY GKEYVTPGSF DFLHKQKFWD LVAKRGDQDS IPKEYAKKLI
        msa25122.2{100_2603}  LLFSSQYFQY GKEYVTPGSF DFLHKQKFWD LVAKRGDQDS IPKEYAKKLI
        msa25122.2{100_A909}  LLFSSQYFQY GKEYVTPGSF DFLHKQKFWD LVAKRGDQDS IPKEYAKKLI
      msa25122.2{100_CJB110}  LLFSSQYFQY GKEYVTPGSF DFLHKQKFWD LVAKRGDQDS IPKEYAKKLI
        msa25122.2{100_COH1}  LLFSSQYFQY GKEYVTPGSF DFLHKQKFWD LVAKRGDQDS IPKEYAKKLI
   msa25122.2{100_JM9130013}  LLFSSQYFQY GKEYVTPGSF DFLHKQKFWD LVAKRGDQDS IPKEYAKKLI
        msa25122.2{100_M732}  LLFSSQYFQY GKEYVTPGSF DFLHKQKFWD LVAKRGDQDS IPKEYAKKLI
        msa25122.2{100_M781}  LLFSSQYFQY GKEYVTPGSF DFLHKQKFWD LVAKRGDQDS IPKEYAKKLI
                  Consensus  ********** ********** ********** ********** **********

                             101                                                 150
        msa25122.2{100_090}  AMHQKRGDKI VFITGRTRGS MYKEGEVDKT AKALAKDFKL DKPIAVNYTG
      msa25122.2{100_18RS21}  AMHQKRGDKI VFITGRTRGS MYKEGEVDKT AKALAKDFKL DKPIAVNYTG
        msa25122.2{100_2603}  AMHQKRGDKI VFITGRTRGS MYKEGEVDKT AKALAKDFKL DKPIAVNYTG
        msa25122.2{100_A909}  AMHQKRGDKI VFITGRTRGS MYKEGEVDKT AKALAKDFKL DKPIAVNYTG
      msa25122.2{100_CJB110}  AMHQKRGDKI VFITGRTRGS MYKEGEVDKT AKALAKDFKL DKPIAVNYTG
        msa25122.2{100_COH1}  AMHQKRGDKI VFITGRTRGS MYKEGEVDKT AKALAKDFKL DKPIAVNYTG
   msa25122.2{100_JM9130013}  AMHQKRGDKI VFITGRTRGS MYKEGEVDKT AKALAKDFKL DKPIAVNYTG
        msa25122.2{100_M732}  AMHQKRGDKI VFITGRTRGS MYKEGEVDKT AKALAKDFKL DKPIAVNYTG
        msa25122.2{100_M781}  AMHQKRGDKI VFITGRTRGS MYKEGEVDKT AKALAKDFKL DKPIAVNYTG
                  Consensus  ********** ********** ********** ********** **********

                             151                                                 200
        msa25122.2{100_090}  DKPKKPYKYD KSYYIKKYGS DIHYGDSDDD IHAAREAGAR PIRILRAPNS
      msa25122.2{100_18RS21}  DKPKKPYKYD KSYYIKKYGS DIHYGDSDDD IHAAREAGAR PIRILRAPNS
        msa25122.2{100_2603}  DKPKKPYKYD KSYYIKKYGS DIHYGDSDDD IHAAREAGAR PIRILRAPNS
        msa25122.2{100_A909}  DKPKKPYKYD KSYYIKKYGS DIHYGDSDDD IHAAREAGAR PIRILRAPNS
      msa25122.2{100_CJB110}  DKPKKPYKYD KSYYIKKYGS DIHYGDSDDD IHAAREAGAR PIRILRAPNS
        msa25122.2{100_COH1}  DKPKKPYKYD KSYYIKKYGS DIHYGDSDDD IHAAREAGAR PIRILRAPNS
   msa25122.2{100_JM9130013}  DKPKKPYKYD KSYYIKKYGS DIHYGDSDDD IHAAREAGAR PIRILRAPNS
        msa25122.2{100_M732}  DKPKKPYKYD KSYYIKKYGS DIHYGDSDDD IHAAREAGAR PIRILRAPNS
        msa25122.2{100_M781}  DKPKKPYKYD KSYYIKKYGS DIHYGDSDDD IHAAREAGAR PIRILRAPNS
                  Consensus  ********** ********** ********** ********** **********

                             201                  221
        msa25122.2{100_090}  TNLPLPEAGG YGEEVLENSA Y
      msa25122.2{100_18RS21}  TNLPLPEAGG YGEEVLENSA Y
        msa25122.2{100_2603}  TNLPLPEAGG YGEEVLENSA Y
        msa25122.2{100_A909}  TNLPLPEAGG YGEEVLENSA Y
      msa25122.2{100_CJB110}  TNLPLPEAGG YGEEVLENSA Y
        msa25122.2{100_COH1}  TNLPLPEAGG YGEEVLENSA Y
   msa25122.2{100_JM9130013}  TNLPLPEAGG YGEEVLENSA Y
        msa25122.2{100_M732}  TNLPLPEAGG YGEEVLENSA Y
        msa25122.2{100_M781}  TNLPLPEAGG YGEEVLENSA Y
                  Consensus  ********** ********** *
```

Table 41: Comparative Sequences relating to SAG0649

SEQ ID NO. 4101: SAG0649 FROM 2603 V/R GBS TYPE V STRAIN
ATGAAAAAGAGACAAAAAATA
TGGAGAGGGTTATCAGTTACTTTACTAATCCTGTCCCAAATTCCATTTGGTATATTGGTA
CAAGGTGAAACCCAAGATACCAATCAAGCACTTGGAAAAGTAATTGTTAAAAAAACGGGA
GACAATGCTACACCATTAGGCAAAGCGACTTTTGTGTTAAAAAATGACAATGATAAGTCA
GAAACAAGTCACGAAACGGTAGAGGGTTCTGGAGAAGCAACCTTTGAAAACATAAAACCT
GGAGACTACACATTAAGAGAAGAAACAGCACCAATTGGTTATAAAAAAACTGATAAAACC
TGGAAAGTTAAAGTTGCAGATAACGGAGCAACAATAATCGAGGGTATGGATGCAGATAAA
GCAGAGAAACGAAAAGAAGTTTTGAATGCCCAATATCCAAAATCAGCTATTTATGAGGAT
ACAAAAGAAAATTACCCATTAGTTAATGTAGAGGGTTCCAAAGTTGGTGAACAATACAAA
GCATTGAATCCAATAAATGGAAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATCA
AAAAAAATTACAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTGAATTAACTGTT
GAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAACCACTAGATGTCGTTGTGCTA
TTAGATAATTCAAATAGTATGAATAATGAAAGAGCCAATAATTCTCAAAGAGCATTAAAA
GCTGGGGAAGCAGTTGAAAAGCTGATTGATAAAATTACATCAAATAAAGACAATAGAGTA
GCTCTTGTGACATATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGGA
GTTGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATTATCATAAAACT
ACTTTTACAGCAACTACACATAATTACAGTTATTTAAATTTAACAAATGATGCTAACGAA
GTTAATATTCTAAAGTCAAGAATTCCAAAGGAAGCGGAGCATATAAATGGGGATCGCACG
CTCTATCAATTTGGTGCGACATTTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTA
GAGACACAAAGTTCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCCT
ACGATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACCAAAACCAGTTT
AATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATTCTCCAAGAGGATTTTATAATC
AATGGTGATGATTATCAAATAGTAAAAGGAGATGGAGAGAGTTTTAAACTGTTTTCGGAT
AGAAAAGTTCCTGTTACTGGAGGAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAA
CTCTCTGTAATGAGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTATTGG
AGAGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTTCTGCAACGAAA
CAAATCAAAACTCATGGTGAGCCAACAACATTATACTTTAATGGAAATATAAGACCTAAA
GGTTATGACATTTTTACTGTTGGGATTGGTGTAAACGGAGATCCTGGTGCAACTCCTCTT
GAAGCTGAGAAATTTATGCAATCAATATCAAGTAAAACAGAAAATTATACTAATGTTGAT
GATACAAATAAAATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAAA
CATTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTGAATTCCAATTA
AAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTGGTTGGAAATGATGGCAGTCAA
TTAAAAAATGGTGTGGCTCTTGGTGGACCAAACAGTGATGGGGGAATTTTAAAAGATGTT
ACAGTGACTTATGATAAGACATCTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGT
GGACAAAAAGTAGTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAAA
TTTTACAATACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAGAACCAAATACT
ATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGTGAGTTTCCGGTACTAACCATC
AGTAATCAGAAGAAAATGGGTGAGGTTGAATTTATTAAAGTTAATAAAGACAAACATTCA
GAATCGCTTTTGGGAGCTAAGTTTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAG
CAATTTGTTCCAGAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAAA
GCACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATGGCTATATAGAG
GTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAATGGAGAAGTTACGAACCTGAAA
GCAGATCCAAATGCTAATAAAAATCAAATCGGGTATCTTGAAGGAAATGGTAAACATCTT
ATTACCAACACTCCCAAACGCCCACCAGGTGTTTTTCCTAAAACAGGGGGAATTGGTACA
ATTGTCTATATATTAGTTGGTTCTACTTTTATGATACTTACCATTTGTTCTTTCCGTCGT
AAACAATTG

SEQ ID NO. 4102: SAG0649 FROM 090 GBS TYPE Ia STRAIN
GGTGAAACCCAAGATACCAATCAAGCACTTGGAAAAG
TAATTGTTAAAAAAACGGGAGACAATGCTACACCATTAGGCAAAGCGACT
TTTGTGTTAAAAAATGACAATGATAAGTCAGAAACAAGTCACGAAACGGT
AGAGGGTTCTGGAGAAGCAACCTTTGAAAACATAAAACCTGGAGACTACA
CATTAAGAGAAGAAACAGCACCAATTGGTTATAAAAAAACTGATAAAACC
TGGAAAGTTAAAGTTGCAGATAACGGAGCAACAATAATCGAGGGTATGGA
TGCAGATAAAGCAGAGAAACGAAAAGAAGTTTTGAATGCCCAATATCCAA
AATCAGCTATTTATGAGGATACAAAAGAAAATTACCCATTAGTTAATGTA
GAGGGTTCCAAAGTTGGTGAACAATACAAAGCATTGAATCCAATAAATGG
AAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATCAAAAAAAATTA
CAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTGAATTAACTGTT
GAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAACCACTAGATGT
CGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGAAAGAGCCAATA
ATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAAAGCTGATTGAT
AAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTGACATATGCCTC
AACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGGAGTTGCCGATC
AAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATTATCATAAAACT
ACTTTTACAGCAACTACACATAATTACAGTTATTTAAATTTAACAAATGA
TGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAAGGAAGCGGAGC
ATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGACATTTACTCAA
AAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAAAGTTCTAATGC
TAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCCTACGATGTCTT
ATGCCATAAATTTTAATCCTTATATATCAACATCTTACCAAAACCAGTTT
AATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATTCTCCAAGAGGA
TTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGGAGATGGAGAGA
GTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTGGAGGAACGACA
CAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTAATGAGTAATGA
GGGATATGCAATTAATAGTGGATATATTTaTCTCTATTGGAGAGATTACA
ACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTTCTGCAACGAAA
CAAATCAAAACTCATGGTGAGCCAACAACATTATACTTTAATGGAAATAT
AAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGGTGTAAACGGAG
ATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGCAATCAATATCA
AGTAAAACAGAAAATTATACTAATGTTGATGATACAAATAAAATTTATGA
TGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAAACATTCTATTG

Table 41: Comparative Sequences relating to SAG0649

TTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTGAATTCCAATTA
AAAAATGGTCAAAGTTTTACACATGATGATTACGtTTTGGtTGGAAATGA
tGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACCAAACAGTGATG
GGGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGACATCTCAAACC
ATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAAGTAGTTCTTAC
CTATGATGTACGTTTAAAAGATAACTATATAAGTAACAAATTTTACAATA
CAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAGAACCAAATACT
ATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGTGAGTTTCCGGT
ACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGAATTTATTAAAG
TTAATAAAGACAAACATTCAGAATCGCTTTTGGGAGCTAAGTTTCAACTT
CAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTTCCAGAGGGAAG
TGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAAAGCACTTCAAG
ATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATGGCTATATAGAG
GTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAATGGAGAAGTTAC
GAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAATCGGGTATCTTG
AAGGAAATGGTAAACATCTTATTACCAACACTCCCAAACGCCCACCAGGT
GTT

SEQ ID NO. 4103: SAG0649 FROM A909 GBS TYPE 1a STRAIN
GGTGAAACCCAAGATACCAATCAAGCACTTGGAAAA
GTAATTGTTAAAAAAACGGGGGACAATGCTACACCATTAGGCAAAGCGAC
TTTTGTGTTAAAAAATGACAATGATAAGTCAgAAACAAGTCACGAAACGG
TAGAGGGTTCTGGAGAAgCAACCTTTGAAAACATAAAACCTGGAGACTAC
ACATTAAGAGAAGAAACAGCACCAATTGGTTATAAAAAAACTGATAAAAC
CTGGAAAGTTAAAGTTGCAGATAACGGAGCAACAATAATCGAGGGTATGG
ATGCAGATAAAGCAGAGAAACGAAAAGAAGTTTTGAATGCCCAATATCCA
AAATCAGCTATTTATGAGGATACAAAAGAAAATTACCCATTAgTTAATGT
AGAGGGTTCCAAAGTTGGTGAACAATACAAAGCATTGAATCCAATAAATG
GAAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATCAAAAAAAATT
ACAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTGAATTAACTGT
TGAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAACCACTAGATG
TCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGAAAGAGCCAAT
AATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAAAGCTGATTGA
TAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTGACATATGCCT
CAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGGAGTTGCCGAT
CAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATTATCATAAAAC
TACTTTTACAGCAACTACACATAATTACAGTTATTTAAATTTAACAAATG
ATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAAGGAAGCGGAG
CATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGACATTTACTCA
AAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAAAGTTCTAATG
CTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCCTACGATGTCT
TATGCCATAAATTTTAATCCTTATATATCAACATCTTACCAAAACCAGTT
TAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATTCTCCAAGAGG
ATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGGAGATGGAGAG
AGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTGGAGGAACGAC
ACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTAATGAGTAATG
AGGGATATGCAATTAATAGTGGATATATTTATCTCTATTGGAGAGATTAC
AACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTTCTGCAACGAA
ACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTTAATGGAAATA
TAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGGTGTAAACGGA
GATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGCAATCAATATC
AAGTAAAACAGAAAATTATACTAATGTTGATGATACAAATAAAATTTATG
ATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAAACATTCTATT
GTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTGAATTCCAATT
AAAAAATGGTCAAAGTTTTACACATGATGATTACGtTTTGGtTGGAAATG
AtGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACCAAACAGTGAT
GGGGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGACATCTCAAAC
CATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAAGTAGTTCTTA
CCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAAATTTTACAAT
ACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAGAACCAAATAC
TATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGTGAGTTTCCGG
TACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGAATTTATTAAA
GTTAATAAAGACAAACATTCAGAATCGCTTTTGGGAGCTAAGTTTCAACT
TCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTTCCAGAGGGAA
GTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAAAGCACTTCAA
GATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATGGCTATATAGA
GGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAATGGAGAAGTTA
CGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAATCGGGTATCTT
GAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAACGCCCACCAGG
TGTT

SEQ ID NO. 4104: SAG0649 FROM 18RS21 GBS TYPE II STRAIN
GGTGAAACCCAAGATACCAATCAAGCAC
TTGGAAAAGTAATTGTTAAAAAAACGGGAGACAaTGCTACACCaTTAGGC
AAAGCGACTTTTGTGTTAAAAAATGACAATGATAAGTCAGAAACAAGTCA
CGAAACGGTAGAGGGTTCTGGAGAAgCAACCTTTGAAAACATAAAACCTG
GAGACTACACATTAAGAGAAGAAACAGCACCAATTGGTTATAAAAAAACT
GATAAAACCTGGAAAGTTAAAGTTGCAGATAACGGAGCAACAATAATCGA
GGGTATGGATGCAGATAAAGCAGAGAAACGAAaAGAAGTTTTGAATGCCC
AATATCCAAAATCAGCTATTTATGAGGATACAAAAGAAAATTACCCATTA
GTTAATGTAGAGGGTTCCAAAGTTGGTGAACAATACAAAGCATTGAATCC
AATAAATGGAAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATCAA
AAAAAATTaCaGGGGTCAATGATCTCGATAAGAATAAATATAAAATTGAA

Table 41: Comparative Sequences relating to SAG0649

```
TTAACTGTTGAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAACC
ACTAGATGTCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGAAA
GAGCCAATAATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAAAG
CTGATTGATAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTGAC
ATATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGGAG
TTGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATTAT
CATAAAACTACTTTTACAGCAACTACACATAATTACAGTTATTTAAATTT
AACAAATGATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAAGG
AAGCGGAGCATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGACA
TTTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAAAG
TTCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCCTA
CGATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACCAA
AACCAGTTTAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATTCT
CCAAGAGGATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGGAG
ATGGAGAGAGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTGGA
GGAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTAAT
GAGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTATTGGA
GAGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTTCT
GCAACGAAACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTTAA
TGGAAATATAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGGTG
TAAACGGAGATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGCAA
TCAATATCAAGTAAAACAGAAAATTATACTAATGTTGATGATACAAATAA
AATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAAAC
ATTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTGAA
TTCCAATTAAAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTGGT
TGGAAATGATGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACCAA
ACAGTGATGGGGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGACA
TCTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAAGT
AGTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAAAT
TTTACAATACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAGAA
CCAAATACTATTtcGtgATTtCCCAATTCCCAAAATTCGTGATGTTCGTGA
GTTTCCGGTACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGAAT
TTATTAAAGTTAATAAAGACAAACATTCAGAATCGCTTTTGGGAGCTAAG
TTTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTTCC
AGAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAAAG
CACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATGGC
TATATAGAGGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAATGG
AGAAGTTACGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAATCG
GGTATCTTGAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAACGC
CCACCAGGTGTT
```

SEQ ID NO. 4105: SAG0649 FROM M732 GBS TYPE III STRAIN

```
GGTGAAACCCAAGATACCAATCAAGCACT
TGGAAAAGTAATTGTTAAAAAAACGGGAGACAaTGCTACACCATTAGGCA
AAGCGACTTTTGTGTTAAAAAATGACAATGATAAGTCAGAAACAAGTCAC
GAAACGGTAGAGGGTTCTGGAGAAGCAACCTTTGAAAACATAAAACCTGG
AGACTACACATTAAGAGAAGAAACAGCACCAATTGGTTATAAAAAAACTG
ATAAAACCTGGAAAGTTAAAGTTGCAGATAACGGAGCAACAATAATCGAG
GGTATGGATGCAGATAAAGCAGAGAAACGAAAAGAAGTTTTGAATGCCCA
ATATCCAAAATCAGCTATTTATGAGGATACAAAAGAAAATTACCCATTAg
TTAATGTAGAGGGTTCCAAAGTTGGTGAACAATACAAAGCATTGAATCCA
ATAAATGGAAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATCAAA
AAAAAaTaCaGGGGTCAATGATCTCGATAAGAATAAATATAAAATTGAAT
TAACTGTTGAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAACCA
CTAGATGTCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGAAAG
AGCCAATAATTCTCAAAGAGCATTAAAaGCTGGGGAAGCAGTTGAAAAGC
TGATTGATAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTGACA
TATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGGAGT
TGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATTATC
ATAAAACTACTTTTACAGCAACTACACATAATTACAGTTATTTAAATTTA
ACAAATGATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAAGGA
AGCGGAGCATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGACAT
TTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAAAGT
TCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCCTAC
GATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACCAAA
ACCAGTTTAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATTCTC
CAAGAGGATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGGAGA
TGGAGAGAGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTGGAG
GAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTAATG
AGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTATTGGAG
AGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTTCTG
CAACGAAACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTTAAT
GGAAATATAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGGTGT
AAACGGAGATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGCAAT
CAATATCAAGTAAAACAGAAAATTATACTAATGTTGATGATACAAATAAA
ATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAAACA
TTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTGAAT
TCCAATTAAAAAATGGTCAAAGTTTTACACATGATGATTACGtTTTGGtT
GGAAATGAtGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACCAAA
CAGTGATGGGGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGACAT
CTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAAGTA
GTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAAATT
TTACAATACAAATAATCGTACAaCGCTAAGTCCGAAGAGTGAAAAAGAAC
```

Table 41: Comparative Sequences relating to SAG0649

CAAATACTATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGTGAG
TTTCCGGTACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGAATT
TATTAAAGTTAATAAAGACAAACATTCAGAATCGCTTTTGGGAGCTAAGT
TTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTTCCA
GAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAAAGC
ACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATGGCT
ATATAGAGGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAATGGA
GAAGTTACGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAATCGG
GTATCTTGAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAACGCC
CACCAGGTGTT

SEQ ID NO. 4106: SAG0649 FROM COH1 GBS TYPE III STRAIN
GGTGAAACCCAAGATACCAATCAAGCACTTGGAAAAG
TAATTGTTAAAAAAACGGGAGACAaTGCTACACCATTAGGCAAAGCGACT
TTTGTGTTAAAAAATGACAATGATAAGTCAGAAACAAGTCACGAAACGGT
AGAGGGTTCTGGArAAGCAACCTTTGAAAACATAAAACCTGGAGACTACA
CATTAAGAGAAGAAACAGCACCAATTGGTTATAAAAAAACTGATAAAACC
TGGAAAGTTAAAGTTGCAGATAACGGAGCAACAATAATCGAGGGTATGGA
TGCAGATAAAGCAGAGAAACGAAAAGAAGTTTTGAATGCCCAATATCCAA
AATCAGCTATTTATGAGGATACAAAAGAAAATTACCCATTAgTTAATGTA
GAGGGTTCCAAAGTTGGTGAACAATaCAAAGCATTGAATCCAATAAATGG
AAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATCAAAAAAAAATA
CAGGGGTCAATGATCTCgATAAGAATAAATATAAAATTGAATTAACTGTT
GAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAACCACTAGATGT
CGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGAAAGAGCCAATA
ATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAAAGCTGATTGAT
AAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTGACATATGCCTC
AACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGGAGTTGCCGATC
AAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATTATCATAAAACT
ACTTTTACAGCAACTACACATAATTACAGTTATTTAAATTTAACAAATGA
TGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAAGGAAGCGGAGC
ATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGACATTTACTCAA
AAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAAAGTTCTAATGC
TAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCCTACGATGTCTT
ATGCCATAAATTTTAATCCTTATATATCAACATCTTACCAAAACCAGTTT
AATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATTCTCCAAGAGGA
TTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGGAGATGGAGAGA
GTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTGGAGGAACGACA
CAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTAATGAGTAATGA
GGGATATGCAATTAATAGTGGATATATTTATCTCTATTGGAGAGATTACA
ACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTTCTGCAACGAAA
CAAATCAAAACTCATGGTGAGCCAACAACATTATACTTTAATGGAAATAT
AAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGGTGTAAACGGAG
ATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGCAATCAATATCA
AGTAAAACAGAAAATTATACTAATGTTGATGATACAAATAAAATTTATGA
TGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAAACATTCTATTG
TTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTGAATTCCAATTA
AAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTGGTTGGAAATGA
TGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACCAAACAGTGATG
GGGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGACATCTCAAACC
ATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAAGTAGTTCTTAC
CTATGATGTACGTTTAAAAGATAACTATATAAGTAACAAATTTTACAATA
CAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAGAACCAAATACT
ATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGTGAGTTTCCGGT
ACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGAATTTATTAAAG
TTAATAAAGACAAACATTCAgAATCGCTTTTGGGAGCTAAGTTTCAACTT
CAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTTCCAGAGGGAAG
TGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAAAGCACTTCAAG
ATGGTAACTATAAATTATATGAAATTTCAAGTCCAgATGGCTATATAGAG
GTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAATGGAGAAGTTAC
GAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAATCGGGTATCTTG
AAGGAAATGGTAAACATCTTATTACCAACACTCCCAAACGCCCACCAGGT
GTT

SEQ ID NO. 4107: SAG0649 FROM M781 GBS TYPE III STRAIN
TTGGAAAAGTAATTGTTAAAAAAACGGGAGACACTGCTACACCATTAGGC
AAAGCGACTTTTGTGTTAAAAAATGACAATGATAAGTCAGAAACAAGTCA
CGAAACGGTAGAGGGTTCTGGAAAAGCAACCTTTGAAAACATAAAACCTG
GAGACTACACATTAAGAGAAGAAACAGCACCAATTGGTTATAAAAAAACT
GATAAAACCTGGAAAGTTAAAGTTGCAGATAACGGAGCAmCAATAATCGA
GGGTATGGATGCAGATAAAGCAGAGAAACGAAAAGAAGTTTTGAATGCCC
AATATCCAAAATCAGCTATTTATGAGGATACAAAAGAAAATTACCCATTA
gTTAATGTAGAGGGTTCCAAAGTTGGTGAACAATACAAAGCATTGAATCC
AATAAATGGAAAAGATGGTCgAAGAGAGATTGCTGAAGGTTGGTTATCAA
AAAAAATTACaGGGGTCAATGATCTCGATAAGAATAAATATAAAATTGAA
TTAACTGTTGAGGGTAAAACCACTGTTGAAACgAAAGAACTTAATCAACC
ACTAGATGTCGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGAAA
GAGCCAATAATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAAAG
CTGATTGATAAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTGAC
ATATGCCTCAACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGGAG
TTGCCGATCAAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATTAT
CATAAAACTACTTTTACAGCAACTACACATAATTACAGTTATTTAAATTT
AACAAATGATGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAAGG

Table 41: Comparative Sequences relating to SAG0649

AAGCGGAGCATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGACA
TTTACTCAAAAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAAAG
TTCTAATGCTAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCCTA
CGATGTCTTATGCCATAAATTTTAATCCTTATATATCAACATCTTACCAA
AACCAGTTTAATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATTCT
CCAAGAGGATTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGGAG
ATGGAGAGAGTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTGGA
GGAACGACACAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTAAT
GAGTAATGAGGGATATGCAATTAATAGTGGATATATTTATCTCTAtTGGA
GAGATTACAACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTTCT
GCAACGAAACAAATCAAAACTCATGGTGAGCCAACAACATTATACTTTAA
TGGAAATATAAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGGTG
TAAACGGAGATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGCAA
TCAATATCAAGTAAAACAGAAAATTATACTAATGTTGATGATACAAATAA
AATTTATGATGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAAAC
ATTCTATTGTTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTGAA
TTCCAATTAAAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTGGT
TGGAAATGATGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACCAA
ACAGTGATGGGGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGACA
TCTCAAACCATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAAGT
AGTTCTTACCTATGATGTACGTTTAAAAGATAACTATATAAGTAACAAAT
TTTACAATACAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAGAA
CCAAATACTATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGTGA
GTTTCCGGTACTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGAAT
TTATTAAAGTTAATAAAGACAAACATTCAGAATCGCTTTTGGGAGCTAAG
TTTCAACTTCAGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTTCC
AGAGGGAAGTGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAAAG
CACTTCAAGATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATGGC
TATATAGAGGTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAATGG
AGAAGTTACGAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAATCG
GGTATCTTGAAGGAAATGGTAAACATCTTATTACCAACACTCCCAAACGC
CCACCAGGTGTT

SEQ ID NO. 4108: SAG0649 FROM CJB GBS NONTYPEABLE STRAIN
GGTGAAACCCAAGATACCAATCAAGCACTTGGAAAAGT
AATTGTTAAAAAAACGGGAGACAaTGCTACACCATTAGGCAAAGCGACTT
TTGTGTTAAAAAATGACAATGATAAGTCAGAAACAAGTCACGAAACGGTA
GAGGGTTCTGGArAAGCAACCTTTGAAAACATAAAACCTGGAGACTACAC
ATTAAGAGAAGAAACAGCACCAATTGGTTATAAAAAAACTGATAAAACCT
GGAAAGTTAAAGTTGCAGATAACGGAGCAACAATAATCGAGGGTATGGAT
GCAGATAAAGCAGAGAAACGAAAAGAAGTTTTGAATGCCCAATATCCAAA
ATCAGCTATTTATGAGGATACAAAAGAAAATTACCCATTAgTTAATGTAG
AGGGTTCCAAAGTTGGTGAACAATACAAAGCATTGAATCCAATAAATGGA
AAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATCAAAAAAAATTAC
aGGGGTCAATGATCTCGATAAGAATAAATATAAAATTGAATTAACTGTTG
AGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAACCACTAGATGTC
GTTGTGCTATTAgATAATTCAAATAGTATGAATAATGAAAGAGCCAATAA
TTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAAAGCTGATTGATA
AAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTGACATATGCCTCA
ACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGGAGTTGCCGATCA
AAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATTATCATAAAACTA
CTTTTACAGCAACTACACATAATTACAGTTATTTAAATTTAACAAATGAT
GCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAAGGAAGCGGAGCA
TATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGACATTTACTCAAA
AAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAAAGTTCTAATGCT
AGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCCTACGATGTCTTA
TGCCATAAATTTTAATCCTTATATATCAACATCTTACCAAAACCAGTTTA
ATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATTCTCCAAGAGGAT
TTTATAATCAATGGTGATGATTATCAAATAGTAAAAGGAGATGGAGAGAG
TTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTGGAGGAACGACAC
AAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTAATGAGTAATGAG
GGATATGCAATTAATAGTGGATATATTTATCTCTATTGGAGAGATTACAA
CTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTTCTGCAACGAAAC
AAATCAAAACTCATGGTGAGCCAACAACATTATACTTTAATGGAAATATA
AGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGGTGTAAACGGAGA
TCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGCAATCAATATCAA
GTAAAACAGAAAATTATACTAATGTTGATGATACAAATAAAATTTATGAT
GAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAAACATTCTATTGT
TGATGGAAATGTGACTGATCCTATGGGAGAGATGATTGAATTCCAATTAA
AAAATGGTCAAAGTTTTACACATGATGATTACGTTTTGGTTGGAAATGAt
GGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACCAAACAGTGATGG
GGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGACATCTCAAACCA
TCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAAGTAGTTCTTACC
TATGATGTACGTTTAAAAGATAACTATATAAGTAACAAATTTTACAATAC
AAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAGAACCAAATACTA
TTCGTGATTTCCCAATtCCCAAAATTCGTGATGTTCGTGAGTTTCCGGTA
CTAACCATCAGTAATCAGAAGAAAATGGGTGAGGTTGAATTTATTAAAGT
TAATAAAGACAAACATTCAGAATCGCTTTTGGGAGCTAAGTTTCAACTTC
AGATAGAAAAAGATTTTTCTGGGTATAAGCAATTTGTTCCAGAGGGAAGT
GATGTTACAACAAAGAATGATGGTAAAATTTATTTTAAAGCACTTCAAGA
TGGTAACTATAAATTATATGAAATTTCAAGTCCAGATGGCTATATAGAGG
TTAAAACGAAACCTGTTGTGACATTTACAATTCAaAATGGAGAAGTTACG
AACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAATCGGGTATCTTGA

Table 41: Comparative Sequences relating to SAG0649

AGGAAATGGTAAACATCTTATTACCAACACTCCCAAACGCCCACCAGGTG
TT

SEQ ID NO. 4109: SAG0649 FROM JM9130013 GBS TYPE VIII STRAIN
GGTGAAACCCAAGATACCAATCAAGCACTTGGAAAAG
TAATTGTTAAAAAAACGGGAGACAATGCTACACCATTAGGCAAAGCGACT
TTTGTGTTAAAAAATGACAATGATAAGTCAGAAACAAGTCACGAAACGGT
AGAGGGTTCTGGAGAAGCAACCTTTGAAAACATAAAACCTGGAGACTACA
CATTAAGAGAAGAAACAGCACCAATTGGTTATAAAAAAACTGATAAAACC
TGGAAAGTTAAAGTTGCAGATAACGGAGCAACAATAATCGAGGGTATGGA
TGCAGATAAAGCAGAGAAACGAAAAGAAGTTTTGAATGCCCAATATCCAA
AATCAGCTATTTATGAGGATACAAAAGAAAATTACCCATTAGTTAATGTA
GAGGGTTCCAAAGTTGGTGAACAATACAAAGCATTGAATCCAATAAATGG
AAAAGATGGTCGAAGAGAGATTGCTGAAGGTTGGTTATCAAAAAAAATTA
CAGGGGTCAATGATCTCGATAAGAATAAATATAAAATTGAATTAACTGTT
GAGGGTAAAACCACTGTTGAAACGAAAGAACTTAATCAACCACTAGATGT
CGTTGTGCTATTAGATAATTCAAATAGTATGAATAATGAAAGAGCCAATA
ATTCTCAAAGAGCATTAAAAGCTGGGGAAGCAGTTGAAAAGCTGATTGAT
AAAATTACATCAAATAAAGACAATAGAGTAGCTCTTGTGACATATGCCTC
AACCATTTTTGATGGTACTGAAGCGACCGTATCAAAGGGAGTTGCCGATC
AAAATGGTAAAGCGCTGAATGATAGTGTATCATGGGATTATCATAAAACT
ACTTTTACAGCAACTACACATAATTACAGTTATTTAAATTTAACAAATGA
TGCTAACGAAGTTAATATTCTAAAGTCAAGAATTCCAAAGGAAGCGGAGC
ATATAAATGGGGATCGCACGCTCTATCAATTTGGTGCGACATTTACTCAA
AAAGCTCTAATGAAAGCAAATGAAATTTTAGAGACACAAAGTTCTAATGC
TAGAAAAAAACTTATTTTTCACGTAACTGATGGTGTCCCTACGATGTCTT
ATGCCATAAATTTTAATCCTTATATATCAACATCTTACCAAAACCAGTTT
AATTCTTTTTTAAATAAAATACCAGATAGAAGTGGTATTCTCCAAGAGGA
TTTTATAATCAATGGTGATGATTATCAAATAGTAAAAGGAGATGGAGAGA
GTTTTAAACTGTTTTCGGATAGAAAAGTTCCTGTTACTGGAGGAACGACA
CAAGCAGCTTATCGAGTACCGCAAAATCAACTCTCTGTAATGAGTAATGA
GGGATATGCAATTAATAGTGGATATATTTATCTCTATTGGAGAGATTACA
ACTGGGTCTATCCATTTGATCCTAAGACAAAGAAAGTTTCTGCAACGAAA
CAAATCAAAACTCATGGTGAGCCAACAACATTATACTTTAATGGAAATAT
AAGACCTAAAGGTTATGACATTTTTACTGTTGGGATTGGTGTAAACGGAG
ATCCTGGTGCAACTCCTCTTGAAGCTGAGAAATTTATGCAATCAATATCA
AGTAAAACAGAAAATTATACTAATGTTGATGATACAAATAAAATTTATGA
TGAGCTAAATAAATACTTTAAAACAATTGTTGAGGAAAAACATTCTATTG
TTGATGGAAATGTGACTGATCCTATGGGAGAGATGATTGAATTCCAATTA
AAAAATGGTCAAAGTTTTACACATGATGATTACGTTTTGGTTGGAAATGA
TGGCAGTCAATTAAAAAATGGTGTGGCTCTTGGTGGACCAAACAGTGATG
GGGGAATTTTAAAAGATGTTACAGTGACTTATGATAAGACATCTCAAACC
ATCAAAATCAATCATTTGAACTTAGGAAGTGGACAAAAAGTAGTTCTTAC
CTATGATGTACGTTTAAAAGATAACTATATAAGTAACAAATTTTACAATA
CAAATAATCGTACAACGCTAAGTCCGAAGAGTGAAAAAGAACCAAATACT
ATTCGTGATTTCCCAATTCCCAAAATTCGTGATGTTCGTGAGTTTCCGGT
ACTAACCATCAGTAATCAAAAGAAAATGGGTGAGGTTGAATTTATTAAAG
TTAATAAAGACAAACATTCAGAATCGCTTTTGGGAGCTAAGTTTCAACTT
CAGATAAAAAAAGATTTTTCTGGGTATAAGCAATTTGTTCCAGAGGGAAG
TGATGTTACAACAAAGAATGATGGTAAAATTTATTTTAAAGCACTTCAAG
ATGGTAACTATAAATTATATGAAATTTCAAGTCCAGATGGCTATATAGAG
GTTAAAACGAAACCTGTTGTGACATTTACAATTCAAAATGGAGAAGTTAC
GAACCTGAAAGCAGATCCAAATGCTAATAAAAATCAAATCGGGTATCTTG
AA

Table 41: Comparative Sequences relating to SAG0649

```
PRETTY of: /biotmp/msa178297.2{*}   May 12, 2003 09:22  ..

                              1                                                   50
      msa178297.2{104_090}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    msa178297.2{104_18RS21}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     msa178297.2{104_2603}    atgaaaaaga gacaaaaaat atggagaggg ttatcagtta ctttactaat
    msa178297.2{104_CJB110}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     msa178297.2{104_COH1}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     msa178297.2{104_M732}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     msa178297.2{104_A909}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     msa178297.2{104_M781}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
msa178297.2{104_JM9130013}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
                Consensus     ********** ********** ********** ********** **********

                              51                                                 100
      msa178297.2{104_090}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~ggtgaa acccaagata
    msa178297.2{104_18RS21}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~ggtgaa acccaagata
     msa178297.2{104_2603}    cctgtcccaa attccatttg gtatattggt acaaggtgaa acccaagata
    msa178297.2{104_CJB110}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~ggtgaa acccaagata
     msa178297.2{104_COH1}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~ggtgaa acccaagata
     msa178297.2{104_M732}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~ggtgaa acccaagata
     msa178297.2{104_A909}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~ggtgaa acccaagata
     msa178297.2{104_M781}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
msa178297.2{104_JM9130013}    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~ggtgaa acccaagata
                Consensus     ********** ********** ********** ****------ ----------

                              101                                                150
      msa178297.2{104_090}    ccaatcaagc acTTGGAAAA GTAATTGTTA AAAAAACGGG aGACAaTGCT
    msa178297.2{104_18RS21}   ccaatcaagc acTTGGAAAA GTAATTGTTA AAAAAACGGG aGACAaTGCT
     msa178297.2{104_2603}    ccaatcaagc acTTGGAAAA GTAATTGTTA AAAAAACGGG aGACAaTGCT
    msa178297.2{104_CJB110}   ccaatcaagc acTTGGAAAA GTAATTGTTA AAAAAACGGG aGACAaTGCT
     msa178297.2{104_COH1}    ccaatcaagc acTTGGAAAA GTAATTGTTA AAAAAACGGG aGACAaTGCT
     msa178297.2{104_M732}    ccaatcaagc acTTGGAAAA GTAATTGTTA AAAAAACGGG aGACAaTGCT
     msa178297.2{104_A909}    ccaatcaagc acTTGGAAAA GTAATTGTTA AAAAAACGGG gGACAaTGCT
     msa178297.2{104_M781}    ~~~~~~~~~~ ~~TTGGAAAA GTAATTGTTA AAAAAACGGG aGACAcTGCT
msa178297.2{104_JM9130013}    ccaatcaagc acTTGGAAAA GTAATTGTTA AAAAAACGGG aGACAaTGCT
                Consensus     ---------- --******** ********** ********** -****-****

                              151                                                200
      msa178297.2{104_090}    ACACCATTAG GCAAAGCGAC TTTTGTGTTA AAAAATGACA ATGATAAGTC
    msa178297.2{104_18RS21}   ACACCATTAG GCAAAGCGAC TTTTGTGTTA AAAAATGACA ATGATAAGTC
     msa178297.2{104_2603}    ACACCATTAG GCAAAGCGAC TTTTGTGTTA AAAAATGACA ATGATAAGTC
    msa178297.2{104_CJB110}   ACACCATTAG GCAAAGCGAC TTTTGTGTTA AAAAATGACA ATGATAAGTC
     msa178297.2{104_COH1}    ACACCATTAG GCAAAGCGAC TTTTGTGTTA AAAAATGACA ATGATAAGTC
     msa178297.2{104_M732}    ACACCATTAG GCAAAGCGAC TTTTGTGTTA AAAAATGACA ATGATAAGTC
     msa178297.2{104_A909}    ACACCATTAG GCAAAGCGAC TTTTGTGTTA AAAAATGACA ATGATAAGTC
     msa178297.2{104_M781}    ACACCATTAG GCAAAGCGAC TTTTGTGTTA AAAAATGACA ATGATAAGTC
msa178297.2{104_JM9130013}    ACACCATTAG GCAAAGCGAC TTTTGTGTTA AAAAATGACA ATGATAAGTC
                Consensus     ********** ********** ********** ********** **********

                              201                                                250
      msa178297.2{104_090}    AGAAACAAGT CACGAAACGG TAGAGGGTTC TGGAgAAGCA ACCTTTGAAA
    msa178297.2{104_18RS21}   AGAAACAAGT CACGAAACGG TAGAGGGTTC TGGAgAAGCA ACCTTTGAAA
     msa178297.2{104_2603}    AGAAACAAGT CACGAAACGG TAGAGGGTTC TGGAgAAGCA ACCTTTGAAA
    msa178297.2{104_CJB110}   AGAAACAAGT CACGAAACGG TAGAGGGTTC TGGArAAGCA ACCTTTGAAA
     msa178297.2{104_COH1}    AGAAACAAGT CACGAAACGG TAGAGGGTTC TGGArAAGCA ACCTTTGAAA
     msa178297.2{104_M732}    AGAAACAAGT CACGAAACGG TAGAGGGTTC TGGAgAAGCA ACCTTTGAAA
     msa178297.2{104_A909}    AGAAACAAGT CACGAAACGG TAGAGGGTTC TGGAgAAGCA ACCTTTGAAA
     msa178297.2{104_M781}    AGAAACAAGT CACGAAACGG TAGAGGGTTC TGGAaAAGCA ACCTTTGAAA
msa178297.2{104_JM9130013}    AGAAACAAGT CACGAAACGG TAGAGGGTTC TGGAgAAGCA ACCTTTGAAA
                Consensus     ********** ********** ********** ****-***** **********

                              251                                                300
      msa178297.2{104_090}    ACATAAAACC TGGAGACTAC ACATTAAGAG AAGAAACAGC ACCAATTGGT
    msa178297.2{104_18RS21}   ACATAAAACC TGGAGACTAC ACATTAAGAG AAGAAACAGC ACCAATTGGT
     msa178297.2{104_2603}    ACATAAAACC TGGAGACTAC ACATTAAGAG AAGAAACAGC ACCAATTGGT
    msa178297.2{104_CJB110}   ACATAAAACC TGGAGACTAC ACATTAAGAG AAGAAACAGC ACCAATTGGT
     msa178297.2{104_COH1}    ACATAAAACC TGGAGACTAC ACATTAAGAG AAGAAACAGC ACCAATTGGT
     msa178297.2{104_M732}    ACATAAAACC TGGAGACTAC ACATTAAGAG AAGAAACAGC ACCAATTGGT
     msa178297.2{104_A909}    ACATAAAACC TGGAGACTAC ACATTAAGAG AAGAAACAGC ACCAATTGGT
     msa178297.2{104_M781}    ACATAAAACC TGGAGACTAC ACATTAAGAG AAGAAACAGC ACCAATTGGT
msa178297.2{104_JM9130013}    ACATAAAACC TGGAGACTAC ACATTAAGAG AAGAAACAGC ACCAATTGGT
                Consensus     ********** ********** ********** ********** **********

                              301                                                350
      msa178297.2{104_090}    TATAAAAAAA CTGATAAAAC CTGGAAAGTT AAAGTTGCAG ATAACGGAGC
    msa178297.2{104_18RS21}   TATAAAAAAA CTGATAAAAC CTGGAAAGTT AAAGTTGCAG ATAACGGAGC
     msa178297.2{104_2603}    TATAAAAAAA CTGATAAAAC CTGGAAAGTT AAAGTTGCAG ATAACGGAGC
    msa178297.2{104_CJB110}   TATAAAAAAA CTGATAAAAC CTGGAAAGTT AAAGTTGCAG ATAACGGAGC
     msa178297.2{104_COH1}    TATAAAAAAA CTGATAAAAC CTGGAAAGTT AAAGTTGCAG ATAACGGAGC
     msa178297.2{104_M732}    TATAAAAAAA CTGATAAAAC CTGGAAAGTT AAAGTTGCAG ATAACGGAGC
     msa178297.2{104_A909}    TATAAAAAAA CTGATAAAAC CTGGAAAGTT AAAGTTGCAG ATAACGGAGC
     msa178297.2{104_M781}    TATAAAAAAA CTGATAAAAC CTGGAAAGTT AAAGTTGCAG ATAACGGAGC
```

Table 41: Comparative Sequences relating to SAG0649

```
msa178297.2{104_JM9130013}   TATAAAAAAA CTGATAAAAC CTGGAAAGTT AAAGTTGCAG ATAACGGAGC
                 Consensus   ********** ********** ********** ********** **********

                             351                                                400
      msa178297.2{104_090}   AaCAATAATC GAGGGTATGG ATGCAGATAA AGCAGAGAAA CGAAAAGAAG
   msa178297.2{104_18RS21}   AaCAATAATC GAGGGTATGG ATGCAGATAA AGCAGAGAAA CGAAAAGAAG
     msa178297.2{104_2603}   AaCAATAATC GAGGGTATGG ATGCAGATAA AGCAGAGAAA CGAAAAGAAG
   msa178297.2{104_CJB110}   AaCAATAATC GAGGGTATGG ATGCAGATAA AGCAGAGAAA CGAAAAGAAG
     msa178297.2{104_COH1}   AaCAATAATC GAGGGTATGG ATGCAGATAA AGCAGAGAAA CGAAAAGAAG
     msa178297.2{104_M732}   AaCAATAATC GAGGGTATGG ATGCAGATAA AGCAGAGAAA CGAAAAGAAG
     msa178297.2{104_A909}   AaCAATAATC GAGGGTATGG ATGCAGATAA AGCAGAGAAA CGAAAAGAAG
     msa178297.2{104_M781}   AmCAATAATC GAGGGTATGG ATGCAGATAA AGCAGAGAAA CGAAAAGAAG
msa178297.2{104_JM9130013}   AaCAATAATC GAGGGTATGG ATGCAGATAA AGCAGAGAAA CGAAAAGAAG
                 Consensus   *-******** ********** ********** ********** **********

                             401                                                450
      msa178297.2{104_090}   TTTTGAATGC CCAATATCCA AAATCAGCTA TTTATGAGGA TACAAAAGAA
   msa178297.2{104_18RS21}   TTTTGAATGC CCAATATCCA AAATCAGCTA TTTATGAGGA TACAAAAGAA
     msa178297.2{104_2603}   TTTTGAATGC CCAATATCCA AAATCAGCTA TTTATGAGGA TACAAAAGAA
   msa178297.2{104_CJB110}   TTTTGAATGC CCAATATCCA AAATCAGCTA TTTATGAGGA TACAAAAGAA
     msa178297.2{104_COH1}   TTTTGAATGC CCAATATCCA AAATCAGCTA TTTATGAGGA TACAAAAGAA
     msa178297.2{104_M732}   TTTTGAATGC CCAATATCCA AAATCAGCTA TTTATGAGGA TACAAAAGAA
     msa178297.2{104_A909}   TTTTGAATGC CCAATATCCA AAATCAGCTA TTTATGAGGA TACAAAAGAA
     msa178297.2{104_M781}   TTTTGAATGC CCAATATCCA AAATCAGCTA TTTATGAGGA TACAAAAGAA
msa178297.2{104_JM9130013}   TTTTGAATGC CCAATATCCA AAATCAGCTA TTTATGAGGA TACAAAAGAA
                 Consensus   ********** ********** ********** ********** **********

                             451                                                500
      msa178297.2{104_090}   AATTACCCAT TAGTTAATGT AGAGGGTTCC AAAGTTGGTG AACAATACAA
   msa178297.2{104_18RS21}   AATTACCCAT TAGTTAATGT AGAGGGTTCC AAAGTTGGTG AACAATACAA
     msa178297.2{104_2603}   AATTACCCAT TAGTTAATGT AGAGGGTTCC AAAGTTGGTG AACAATACAA
   msa178297.2{104_CJB110}   AATTACCCAT TAGTTAATGT AGAGGGTTCC AAAGTTGGTG AACAATACAA
     msa178297.2{104_COH1}   AATTACCCAT TAGTTAATGT AGAGGGTTCC AAAGTTGGTG AACAATACAA
     msa178297.2{104_M732}   AATTACCCAT TAGTTAATGT AGAGGGTTCC AAAGTTGGTG AACAATACAA
     msa178297.2{104_A909}   AATTACCCAT TAGTTAATGT AGAGGGTTCC AAAGTTGGTG AACAATACAA
     msa178297.2{104_M781}   AATTACCCAT TAGTTAATGT AGAGGGTTCC AAAGTTGGTG AACAATACAA
msa178297.2{104_JM9130013}   AATTACCCAT TAGTTAATGT AGAGGGTTCC AAAGTTGGTG AACAATACAA
                 Consensus   ********** ********** ********** ********** **********

                             501                                                550
      msa178297.2{104_090}   AGCATTGAAT CCAATAAATG GAAAAGATGG TCGAAGAGAG ATTGCTGAAG
   msa178297.2{104_18RS21}   AGCATTGAAT CCAATAAATG GAAAAGATGG TCGAAGAGAG ATTGCTGAAG
     msa178297.2{104_2603}   AGCATTGAAT CCAATAAATG GAAAAGATGG TCGAAGAGAG ATTGCTGAAG
   msa178297.2{104_CJB110}   AGCATTGAAT CCAATAAATG GAAAAGATGG TCGAAGAGAG ATTGCTGAAG
     msa178297.2{104_COH1}   AGCATTGAAT CCAATAAATG GAAAAGATGG TCGAAGAGAG ATTGCTGAAG
     msa178297.2{104_M732}   AGCATTGAAT CCAATAAATG GAAAAGATGG TCGAAGAGAG ATTGCTGAAG
     msa178297.2{104_A909}   AGCATTGAAT CCAATAAATG GAAAAGATGG TCGAAGAGAG ATTGCTGAAG
     msa178297.2{104_M781}   AGCATTGAAT CCAATAAATG GAAAAGATGG TCGAAGAGAG ATTGCTGAAG
msa178297.2{104_JM9130013}   AGCATTGAAT CCAATAAATG GAAAAGATGG TCGAAGAGAG ATTGCTGAAG
                 Consensus   ********** ********** ********** ********** **********

                             551                                                600
      msa178297.2{104_090}   GTTGGTTATC AAAAAAAAtT ACAGGGGTCA ATGATCTCGA TAAGAATAAA
   msa178297.2{104_18RS21}   GTTGGTTATC AAAAAAAAtT ACAGGGGTCA ATGATCTCGA TAAGAATAAA
     msa178297.2{104_2603}   GTTGGTTATC AAAAAAAAtT ACAGGGGTCA ATGATCTCGA TAAGAATAAA
   msa178297.2{104_CJB110}   GTTGGTTATC AAAAAAAAtT ACAGGGGTCA ATGATCTCGA TAAGAATAAA
     msa178297.2{104_COH1}   GTTGGTTATC AAAAAAAAaT ACAGGGGTCA ATGATCTCGA TAAGAATAAA
     msa178297.2{104_M732}   GTTGGTTATC AAAAAAAAaT ACAGGGGTCA ATGATCTCGA TAAGAATAAA
     msa178297.2{104_A909}   GTTGGTTATC AAAAAAAAtT ACAGGGGTCA ATGATCTCGA TAAGAATAAA
     msa178297.2{104_M781}   GTTGGTTATC AAAAAAAAtT ACAGGGGTCA ATGATCTCGA TAAGAATAAA
msa178297.2{104_JM9130013}   GTTGGTTATC AAAAAAAAtT ACAGGGGTCA ATGATCTCGA TAAGAATAAA
                 Consensus   ********** ********-* ********** ********** **********

                             601                                                650
      msa178297.2{104_090}   TATAAAATTG AATTAACTGT TGAGGGTAAA ACCACTGTTG AAACGAAAGA
   msa178297.2{104_18RS21}   TATAAAATTG AATTAACTGT TGAGGGTAAA ACCACTGTTG AAACGAAAGA
     msa178297.2{104_2603}   TATAAAATTG AATTAACTGT TGAGGGTAAA ACCACTGTTG AAACGAAAGA
   msa178297.2{104_CJB110}   TATAAAATTG AATTAACTGT TGAGGGTAAA ACCACTGTTG AAACGAAAGA
     msa178297.2{104_COH1}   TATAAAATTG AATTAACTGT TGAGGGTAAA ACCACTGTTG AAACGAAAGA
     msa178297.2{104_M732}   TATAAAATTG AATTAACTGT TGAGGGTAAA ACCACTGTTG AAACGAAAGA
     msa178297.2{104_A909}   TATAAAATTG AATTAACTGT TGAGGGTAAA ACCACTGTTG AAACGAAAGA
     msa178297.2{104_M781}   TATAAAATTG AATTAACTGT TGAGGGTAAA ACCACTGTTG AAACGAAAGA
msa178297.2{104_JM9130013}   TATAAAATTG AATTAACTGT TGAGGGTAAA ACCACTGTTG AAACGAAAGA
                 Consensus   ********** ********** ********** ********** **********

                             651                                                700
      msa178297.2{104_090}   ACTTAATCAA CCACTAGATG TCGTTGTGCT ATTAGATAAT TCAAATAGTA
   msa178297.2{104_18RS21}   ACTTAATCAA CCACTAGATG TCGTTGTGCT ATTAGATAAT TCAAATAGTA
     msa178297.2{104_2603}   ACTTAATCAA CCACTAGATG TCGTTGTGCT ATTAGATAAT TCAAATAGTA
   msa178297.2{104_CJB110}   ACTTAATCAA CCACTAGATG TCGTTGTGCT ATTAGATAAT TCAAATAGTA
     msa178297.2{104_COH1}   ACTTAATCAA CCACTAGATG TCGTTGTGCT ATTAGATAAT TCAAATAGTA
     msa178297.2{104_M732}   ACTTAATCAA CCACTAGATG TCGTTGTGCT ATTAGATAAT TCAAATAGTA
     msa178297.2{104_A909}   ACTTAATCAA CCACTAGATG TCGTTGTGCT ATTAGATAAT TCAAATAGTA
```

Table 41: Comparative Sequences relating to SAG0649

```
      msa178297.2{104_M781}  ACTTAATCAA CCACTAGATG TCGTTGTGCT ATTAGATAAT TCAAATAGTA
  msa178297.2{104_JM9130013}  ACTTAATCAA CCACTAGATG TCGTTGTGCT ATTAGATAAT TCAAATAGTA
                  Consensus  ********** ********** ********** ********** **********

                             701                                                750
      msa178297.2{104_090}   TGAATAATGA AAGAGCCAAT AATTCTCAAA GAGCATTAAA AGCTGGGGAA
    msa178297.2{104_18RS21}  TGAATAATGA AAGAGCCAAT AATTCTCAAA GAGCATTAAA AGCTGGGGAA
     msa178297.2{104_2603}   TGAATAATGA AAGAGCCAAT AATTCTCAAA GAGCATTAAA AGCTGGGGAA
    msa178297.2{104_CJB110}  TGAATAATGA AAGAGCCAAT AATTCTCAAA GAGCATTAAA AGCTGGGGAA
     msa178297.2{104_COH1}   TGAATAATGA AAGAGCCAAT AATTCTCAAA GAGCATTAAA AGCTGGGGAA
     msa178297.2{104_M732}   TGAATAATGA AAGAGCCAAT AATTCTCAAA GAGCATTAAA AGCTGGGGAA
     msa178297.2{104_A909}   TGAATAATGA AAGAGCCAAT AATTCTCAAA GAGCATTAAA AGCTGGGGAA
     msa178297.2{104_M781}   TGAATAATGA AAGAGCCAAT AATTCTCAAA GAGCATTAAA AGCTGGGGAA
  msa178297.2{104_JM9130013}  TGAATAATGA AAGAGCCAAT AATTCTCAAA GAGCATTAAA AGCTGGGGAA
                  Consensus  ********** ********** ********** ********** **********

                             751                                                800
      msa178297.2{104_090}   GCAGTTGAAA AGCTGATTGA TAAAATTACA TCAAATAAAG ACAATAGAGT
    msa178297.2{104_18RS21}  GCAGTTGAAA AGCTGATTGA TAAAATTACA TCAAATAAAG ACAATAGAGT
     msa178297.2{104_2603}   GCAGTTGAAA AGCTGATTGA TAAAATTACA TCAAATAAAG ACAATAGAGT
    msa178297.2{104_CJB110}  GCAGTTGAAA AGCTGATTGA TAAAATTACA TCAAATAAAG ACAATAGAGT
     msa178297.2{104_COH1}   GCAGTTGAAA AGCTGATTGA TAAAATTACA TCAAATAAAG ACAATAGAGT
     msa178297.2{104_M732}   GCAGTTGAAA AGCTGATTGA TAAAATTACA TCAAATAAAG ACAATAGAGT
     msa178297.2{104_A909}   GCAGTTGAAA AGCTGATTGA TAAAATTACA TCAAATAAAG ACAATAGAGT
     msa178297.2{104_M781}   GCAGTTGAAA AGCTGATTGA TAAAATTACA TCAAATAAAG ACAATAGAGT
  msa178297.2{104_JM9130013}  GCAGTTGAAA AGCTGATTGA TAAAATTACA TCAAATAAAG ACAATAGAGT
                  Consensus  ********** ********** ********** ********** **********

                             801                                                850
      msa178297.2{104_090}   AGCTCTTGTG ACATATGCCT CAACCATTTT TGATGGTACT GAAGCGACCG
    msa178297.2{104_18RS21}  AGCTCTTGTG ACATATGCCT CAACCATTTT TGATGGTACT GAAGCGACCG
     msa178297.2{104_2603}   AGCTCTTGTG ACATATGCCT CAACCATTTT TGATGGTACT GAAGCGACCG
    msa178297.2{104_CJB110}  AGCTCTTGTG ACATATGCCT CAACCATTTT TGATGGTACT GAAGCGACCG
     msa178297.2{104_COH1}   AGCTCTTGTG ACATATGCCT CAACCATTTT TGATGGTACT GAAGCGACCG
     msa178297.2{104_M732}   AGCTCTTGTG ACATATGCCT CAACCATTTT TGATGGTACT GAAGCGACCG
     msa178297.2{104_A909}   AGCTCTTGTG ACATATGCCT CAACCATTTT TGATGGTACT GAAGCGACCG
     msa178297.2{104_M781}   AGCTCTTGTG ACATATGCCT CAACCATTTT TGATGGTACT GAAGCGACCG
  msa178297.2{104_JM9130013}  AGCTCTTGTG ACATATGCCT CAACCATTTT TGATGGTACT GAAGCGACCG
                  Consensus  ********** ********** ********** ********** **********

                             851                                                900
      msa178297.2{104_090}   TATCAAAGGG AGTTGCCGAT CAAAATGGTA AAGCGCTGAA TGATAGTGTA
    msa178297.2{104_18RS21}  TATCAAAGGG AGTTGCCGAT CAAAATGGTA AAGCGCTGAA TGATAGTGTA
     msa178297.2{104_2603}   TATCAAAGGG AGTTGCCGAT CAAAATGGTA AAGCGCTGAA TGATAGTGTA
    msa178297.2{104_CJB110}  TATCAAAGGG AGTTGCCGAT CAAAATGGTA AAGCGCTGAA TGATAGTGTA
     msa178297.2{104_COH1}   TATCAAAGGG AGTTGCCGAT CAAAATGGTA AAGCGCTGAA TGATAGTGTA
     msa178297.2{104_M732}   TATCAAAGGG AGTTGCCGAT CAAAATGGTA AAGCGCTGAA TGATAGTGTA
     msa178297.2{104_A909}   TATCAAAGGG AGTTGCCGAT CAAAATGGTA AAGCGCTGAA TGATAGTGTA
     msa178297.2{104_M781}   TATCAAAGGG AGTTGCCGAT CAAAATGGTA AAGCGCTGAA TGATAGTGTA
  msa178297.2{104_JM9130013}  TATCAAAGGG AGTTGCCGAT CAAAATGGTA AAGCGCTGAA TGATAGTGTA
                  Consensus  ********** ********** ********** ********** **********

                             901                                                950
      msa178297.2{104_090}   TCATGGGATT ATCATAAAAC TACTTTTACA GCAACTACAC ATAATTACAG
    msa178297.2{104_18RS21}  TCATGGGATT ATCATAAAAC TACTTTTACA GCAACTACAC ATAATTACAG
     msa178297.2{104_2603}   TCATGGGATT ATCATAAAAC TACTTTTACA GCAACTACAC ATAATTACAG
    msa178297.2{104_CJB110}  TCATGGGATT ATCATAAAAC TACTTTTACA GCAACTACAC ATAATTACAG
     msa178297.2{104_COH1}   TCATGGGATT ATCATAAAAC TACTTTTACA GCAACTACAC ATAATTACAG
     msa178297.2{104_M732}   TCATGGGATT ATCATAAAAC TACTTTTACA GCAACTACAC ATAATTACAG
     msa178297.2{104_A909}   TCATGGGATT ATCATAAAAC TACTTTTACA GCAACTACAC ATAATTACAG
     msa178297.2{104_M781}   TCATGGGATT ATCATAAAAC TACTTTTACA GCAACTACAC ATAATTACAG
  msa178297.2{104_JM9130013}  TCATGGGATT ATCATAAAAC TACTTTTACA GCAACTACAC ATAATTACAG
                  Consensus  ********** ********** ********** ********** **********

                             951                                               1000
      msa178297.2{104_090}   TTATTTAAAT TTAACAAATG ATGCTAACGA AGTTAATATT CTAAAGTCAA
    msa178297.2{104_18RS21}  TTATTTAAAT TTAACAAATG ATGCTAACGA AGTTAATATT CTAAAGTCAA
     msa178297.2{104_2603}   TTATTTAAAT TTAACAAATG ATGCTAACGA AGTTAATATT CTAAAGTCAA
    msa178297.2{104_CJB110}  TTATTTAAAT TTAACAAATG ATGCTAACGA AGTTAATATT CTAAAGTCAA
     msa178297.2{104_COH1}   TTATTTAAAT TTAACAAATG ATGCTAACGA AGTTAATATT CTAAAGTCAA
     msa178297.2{104_M732}   TTATTTAAAT TTAACAAATG ATGCTAACGA AGTTAATATT CTAAAGTCAA
     msa178297.2{104_A909}   TTATTTAAAT TTAACAAATG ATGCTAACGA AGTTAATATT CTAAAGTCAA
     msa178297.2{104_M781}   TTATTTAAAT TTAACAAATG ATGCTAACGA AGTTAATATT CTAAAGTCAA
  msa178297.2{104_JM9130013}  TTATTTAAAT TTAACAAATG ATGCTAACGA AGTTAATATT CTAAAGTCAA
                  Consensus  ********** ********** ********** ********** **********

                             1001                                              1050
      msa178297.2{104_090}   GAATTCCAAA GGAAGCGGAG CATATAAATG GGGATCGCAC GCTCTATCAA
    msa178297.2{104_18RS21}  GAATTCCAAA GGAAGCGGAG CATATAAATG GGGATCGCAC GCTCTATCAA
     msa178297.2{104_2603}   GAATTCCAAA GGAAGCGGAG CATATAAATG GGGATCGCAC GCTCTATCAA
    msa178297.2{104_CJB110}  GAATTCCAAA GGAAGCGGAG CATATAAATG GGGATCGCAC GCTCTATCAA
     msa178297.2{104_COH1}   GAATTCCAAA GGAAGCGGAG CATATAAATG GGGATCGCAC GCTCTATCAA
     msa178297.2{104_M732}   GAATTCCAAA GGAAGCGGAG CATATAAATG GGGATCGCAC GCTCTATCAA
```

Table 41: Comparative Sequences relating to SAG0649

```
     msa178297.2{104_A909}  GAATTCCAAA GGAAGCGGAG CATATAAATG GGGATCGCAC GCTCTATCAA
     msa178297.2{104_M781}  GAATTCCAAA GGAAGCGGAG CATATAAATG GGGATCGCAC GCTCTATCAA
msa178297.2{104_JM9130013}  GAATTCCAAA GGAAGCGGAG CATATAAATG GGGATCGCAC GCTCTATCAA
                 Consensus  ********** ********** ********** ********** **********

                            1051                                              1100
      msa178297.2{104_090}  TTTGGTGCGA CATTTACTCA AAAAGCTCTA ATGAAAGCAA ATGAAATTTT
   msa178297.2{104_18RS21}  TTTGGTGCGA CATTTACTCA AAAAGCTCTA ATGAAAGCAA ATGAAATTTT
     msa178297.2{104_2603}  TTTGGTGCGA CATTTACTCA AAAAGCTCTA ATGAAAGCAA ATGAAATTTT
   msa178297.2{104_CJB110}  TTTGGTGCGA CATTTACTCA AAAAGCTCTA ATGAAAGCAA ATGAAATTTT
     msa178297.2{104_COH1}  TTTGGTGCGA CATTTACTCA AAAAGCTCTA ATGAAAGCAA ATGAAATTTT
     msa178297.2{104_M732}  TTTGGTGCGA CATTTACTCA AAAAGCTCTA ATGAAAGCAA ATGAAATTTT
     msa178297.2{104_A909}  TTTGGTGCGA CATTTACTCA AAAAGCTCTA ATGAAAGCAA ATGAAATTTT
     msa178297.2{104_M781}  TTTGGTGCGA CATTTACTCA AAAAGCTCTA ATGAAAGCAA ATGAAATTTT
msa178297.2{104_JM9130013}  TTTGGTGCGA CATTTACTCA AAAAGCTCTA ATGAAAGCAA ATGAAATTTT
                 Consensus  ********** ********** ********** ********** **********

                            1101                                              1150
      msa178297.2{104_090}  AGAGACACAA AGTTCTAATG CTAGAAAAAA ACTTATTTTT CACGTAACTG
   msa178297.2{104_18RS21}  AGAGACACAA AGTTCTAATG CTAGAAAAAA ACTTATTTTT CACGTAACTG
     msa178297.2{104_2603}  AGAGACACAA AGTTCTAATG CTAGAAAAAA ACTTATTTTT CACGTAACTG
   msa178297.2{104_CJB110}  AGAGACACAA AGTTCTAATG CTAGAAAAAA ACTTATTTTT CACGTAACTG
     msa178297.2{104_COH1}  AGAGACACAA AGTTCTAATG CTAGAAAAAA ACTTATTTTT CACGTAACTG
     msa178297.2{104_M732}  AGAGACACAA AGTTCTAATG CTAGAAAAAA ACTTATTTTT CACGTAACTG
     msa178297.2{104_A909}  AGAGACACAA AGTTCTAATG CTAGAAAAAA ACTTATTTTT CACGTAACTG
     msa178297.2{104_M781}  AGAGACACAA AGTTCTAATG CTAGAAAAAA ACTTATTTTT CACGTAACTG
msa178297.2{104_JM9130013}  AGAGACACAA AGTTCTAATG CTAGAAAAAA ACTTATTTTT CACGTAACTG
                 Consensus  ********** ********** ********** ********** **********

                            1151                                              1200
      msa178297.2{104_090}  ATGGTGTCCC TACGATGTCT TATGCCATAA ATTTTAATCC TTATATATCA
   msa178297.2{104_18RS21}  ATGGTGTCCC TACGATGTCT TATGCCATAA ATTTTAATCC TTATATATCA
     msa178297.2{104_2603}  ATGGTGTCCC TACGATGTCT TATGCCATAA ATTTTAATCC TTATATATCA
   msa178297.2{104_CJB110}  ATGGTGTCCC TACGATGTCT TATGCCATAA ATTTTAATCC TTATATATCA
     msa178297.2{104_COH1}  ATGGTGTCCC TACGATGTCT TATGCCATAA ATTTTAATCC TTATATATCA
     msa178297.2{104_M732}  ATGGTGTCCC TACGATGTCT TATGCCATAA ATTTTAATCC TTATATATCA
     msa178297.2{104_A909}  ATGGTGTCCC TACGATGTCT TATGCCATAA ATTTTAATCC TTATATATCA
     msa178297.2{104_M781}  ATGGTGTCCC TACGATGTCT TATGCCATAA ATTTTAATCC TTATATATCA
msa178297.2{104_JM9130013}  ATGGTGTCCC TACGATGTCT TATGCCATAA ATTTTAATCC TTATATATCA
                 Consensus  ********** ********** ********** ********** **********

                            1201                                              1250
      msa178297.2{104_090}  ACATCTTACC AAAACCAGTT TAATTCTTTT TTAAATAAAA TACCAGATAG
   msa178297.2{104_18RS21}  ACATCTTACC AAAACCAGTT TAATTCTTTT TTAAATAAAA TACCAGATAG
     msa178297.2{104_2603}  ACATCTTACC AAAACCAGTT TAATTCTTTT TTAAATAAAA TACCAGATAG
   msa178297.2{104_CJB110}  ACATCTTACC AAAACCAGTT TAATTCTTTT TTAAATAAAA TACCAGATAG
     msa178297.2{104_COH1}  ACATCTTACC AAAACCAGTT TAATTCTTTT TTAAATAAAA TACCAGATAG
     msa178297.2{104_M732}  ACATCTTACC AAAACCAGTT TAATTCTTTT TTAAATAAAA TACCAGATAG
     msa178297.2{104_A909}  ACATCTTACC AAAACCAGTT TAATTCTTTT TTAAATAAAA TACCAGATAG
     msa178297.2{104_M781}  ACATCTTACC AAAACCAGTT TAATTCTTTT TTAAATAAAA TACCAGATAG
msa178297.2{104_JM9130013}  ACATCTTACC AAAACCAGTT TAATTCTTTT TTAAATAAAA TACCAGATAG
                 Consensus  ********** ********** ********** ********** **********

                            1251                                              1300
      msa178297.2{104_090}  AAGTGGTATT CTCCAAGAGG ATTTTATAAT CAATGGTGAT GATTATCAAA
   msa178297.2{104_18RS21}  AAGTGGTATT CTCCAAGAGG ATTTTATAAT CAATGGTGAT GATTATCAAA
     msa178297.2{104_2603}  AAGTGGTATT CTCCAAGAGG ATTTTATAAT CAATGGTGAT GATTATCAAA
   msa178297.2{104_CJB110}  AAGTGGTATT CTCCAAGAGG ATTTTATAAT CAATGGTGAT GATTATCAAA
     msa178297.2{104_COH1}  AAGTGGTATT CTCCAAGAGG ATTTTATAAT CAATGGTGAT GATTATCAAA
     msa178297.2{104_M732}  AAGTGGTATT CTCCAAGAGG ATTTTATAAT CAATGGTGAT GATTATCAAA
     msa178297.2{104_A909}  AAGTGGTATT CTCCAAGAGG ATTTTATAAT CAATGGTGAT GATTATCAAA
     msa178297.2{104_M781}  AAGTGGTATT CTCCAAGAGG ATTTTATAAT CAATGGTGAT GATTATCAAA
msa178297.2{104_JM9130013}  AAGTGGTATT CTCCAAGAGG ATTTTATAAT CAATGGTGAT GATTATCAAA
                 Consensus  ********** ********** ********** ********** **********

                            1301                                              1350
      msa178297.2{104_090}  TAGTAAAAGG AGATGGAGAG AGTTTTAAAC TGTTTTCGGA TAGAAAAGTT
   msa178297.2{104_18RS21}  TAGTAAAAGG AGATGGAGAG AGTTTTAAAC TGTTTTCGGA TAGAAAAGTT
     msa178297.2{104_2603}  TAGTAAAAGG AGATGGAGAG AGTTTTAAAC TGTTTTCGGA TAGAAAAGTT
   msa178297.2{104_CJB110}  TAGTAAAAGG AGATGGAGAG AGTTTTAAAC TGTTTTCGGA TAGAAAAGTT
     msa178297.2{104_COH1}  TAGTAAAAGG AGATGGAGAG AGTTTTAAAC TGTTTTCGGA TAGAAAAGTT
     msa178297.2{104_M732}  TAGTAAAAGG AGATGGAGAG AGTTTTAAAC TGTTTTCGGA TAGAAAAGTT
     msa178297.2{104_A909}  TAGTAAAAGG AGATGGAGAG AGTTTTAAAC TGTTTTCGGA TAGAAAAGTT
     msa178297.2{104_M781}  TAGTAAAAGG AGATGGAGAG AGTTTTAAAC TGTTTTCGGA TAGAAAAGTT
msa178297.2{104_JM9130013}  TAGTAAAAGG AGATGGAGAG AGTTTTAAAC TGTTTTCGGA TAGAAAAGTT
                 Consensus  ********** ********** ********** ********** **********

                            1351                                              1400
      msa178297.2{104_090}  CCTGTTACTG GAGGAACGAC ACAAGCAGCT TATCGAGTAC CGCAAAATCA
   msa178297.2{104_18RS21}  CCTGTTACTG GAGGAACGAC ACAAGCAGCT TATCGAGTAC CGCAAAATCA
     msa178297.2{104_2603}  CCTGTTACTG GAGGAACGAC ACAAGCAGCT TATCGAGTAC CGCAAAATCA
   msa178297.2{104_CJB110}  CCTGTTACTG GAGGAACGAC ACAAGCAGCT TATCGAGTAC CGCAAAATCA
     msa178297.2{104_COH1}  CCTGTTACTG GAGGAACGAC ACAAGCAGCT TATCGAGTAC CGCAAAATCA
```

Table 41: Comparative Sequences relating to SAG0649

```
     msa178297.2{104_M732}   CCTGTTACTG GAGGAACGAC ACAAGCAGCT TATCGAGTAC CGCAAAATCA
     msa178297.2{104_A909}   CCTGTTACTG GAGGAACGAC ACAAGCAGCT TATCGAGTAC CGCAAAATCA
     msa178297.2{104_M781}   CCTGTTACTG GAGGAACGAC ACAAGCAGCT TATCGAGTAC CGCAAAATCA
 msa178297.2{104_JM9130013}  CCTGTTACTG GAGGAACGAC ACAAGCAGCT TATCGAGTAC CGCAAAATCA
                  Consensus  ********** ********** ********** ********** **********

                             1401                                              1450
      msa178297.2{104_090}   ACTCTCTGTA ATGAGTAATG AGGGATATGC AATTAATAGT GGATATATTT
    msa178297.2{104_18RS21}  ACTCTCTGTA ATGAGTAATG AGGGATATGC AATTAATAGT GGATATATTT
     msa178297.2{104_2603}   ACTCTCTGTA ATGAGTAATG AGGGATATGC AATTAATAGT GGATATATTT
    msa178297.2{104_CJB110}  ACTCTCTGTA ATGAGTAATG AGGGATATGC AATTAATAGT GGATATATTT
     msa178297.2{104_COH1}   ACTCTCTGTA ATGAGTAATG AGGGATATGC AATTAATAGT GGATATATTT
     msa178297.2{104_M732}   ACTCTCTGTA ATGAGTAATG AGGGATATGC AATTAATAGT GGATATATTT
     msa178297.2{104_A909}   ACTCTCTGTA ATGAGTAATG AGGGATATGC AATTAATAGT GGATATATTT
     msa178297.2{104_M781}   ACTCTCTGTA ATGAGTAATG AGGGATATGC AATTAATAGT GGATATATTT
 msa178297.2{104_JM9130013}  ACTCTCTGTA ATGAGTAATG AGGGATATGC AATTAATAGT GGATATATTT
                  Consensus  ********** ********** ********** ********** **********

                             1451                                              1500
      msa178297.2{104_090}   ATCTCTATTG GAGAGATTAC AACTGGGTCT ATCCATTTGA TCCTAAGACA
    msa178297.2{104_18RS21}  ATCTCTATTG GAGAGATTAC AACTGGGTCT ATCCATTTGA TCCTAAGACA
     msa178297.2{104_2603}   ATCTCTATTG GAGAGATTAC AACTGGGTCT ATCCATTTGA TCCTAAGACA
    msa178297.2{104_CJB110}  ATCTCTATTG GAGAGATTAC AACTGGGTCT ATCCATTTGA TCCTAAGACA
     msa178297.2{104_COH1}   ATCTCTATTG GAGAGATTAC AACTGGGTCT ATCCATTTGA TCCTAAGACA
     msa178297.2{104_M732}   ATCTCTATTG GAGAGATTAC AACTGGGTCT ATCCATTTGA TCCTAAGACA
     msa178297.2{104_A909}   ATCTCTATTG GAGAGATTAC AACTGGGTCT ATCCATTTGA TCCTAAGACA
     msa178297.2{104_M781}   ATCTCTATTG GAGAGATTAC AACTGGGTCT ATCCATTTGA TCCTAAGACA
 msa178297.2{104_JM9130013}  ATCTCTATTG GAGAGATTAC AACTGGGTCT ATCCATTTGA TCCTAAGACA
                  Consensus  ********** ********** ********** ********** **********

                             1501                                              1550
      msa178297.2{104_090}   AAGAAAGTTT CTGCAACGAA ACAAATCAAA ACTCATGGTG AGCCAACAAC
    msa178297.2{104_18RS21}  AAGAAAGTTT CTGCAACGAA ACAAATCAAA ACTCATGGTG AGCCAACAAC
     msa178297.2{104_2603}   AAGAAAGTTT CTGCAACGAA ACAAATCAAA ACTCATGGTG AGCCAACAAC
    msa178297.2{104_CJB110}  AAGAAAGTTT CTGCAACGAA ACAAATCAAA ACTCATGGTG AGCCAACAAC
     msa178297.2{104_COH1}   AAGAAAGTTT CTGCAACGAA ACAAATCAAA ACTCATGGTG AGCCAACAAC
     msa178297.2{104_M732}   AAGAAAGTTT CTGCAACGAA ACAAATCAAA ACTCATGGTG AGCCAACAAC
     msa178297.2{104_A909}   AAGAAAGTTT CTGCAACGAA ACAAATCAAA ACTCATGGTG AGCCAACAAC
     msa178297.2{104_M781}   AAGAAAGTTT CTGCAACGAA ACAAATCAAA ACTCATGGTG AGCCAACAAC
 msa178297.2{104_JM9130013}  AAGAAAGTTT CTGCAACGAA ACAAATCAAA ACTCATGGTG AGCCAACAAC
                  Consensus  ********** ********** ********** ********** **********

                             1551                                              1600
      msa178297.2{104_090}   ATTATACTTT AATGGAAATA TAAGACCTAA AGGTTATGAC ATTTTTACTG
    msa178297.2{104_18RS21}  ATTATACTTT AATGGAAATA TAAGACCTAA AGGTTATGAC ATTTTTACTG
     msa178297.2{104_2603}   ATTATACTTT AATGGAAATA TAAGACCTAA AGGTTATGAC ATTTTTACTG
    msa178297.2{104_CJB110}  ATTATACTTT AATGGAAATA TAAGACCTAA AGGTTATGAC ATTTTTACTG
     msa178297.2{104_COH1}   ATTATACTTT AATGGAAATA TAAGACCTAA AGGTTATGAC ATTTTTACTG
     msa178297.2{104_M732}   ATTATACTTT AATGGAAATA TAAGACCTAA AGGTTATGAC ATTTTTACTG
     msa178297.2{104_A909}   ATTATACTTT AATGGAAATA TAAGACCTAA AGGTTATGAC ATTTTTACTG
     msa178297.2{104_M781}   ATTATACTTT AATGGAAATA TAAGACCTAA AGGTTATGAC ATTTTTACTG
 msa178297.2{104_JM9130013}  ATTATACTTT AATGGAAATA TAAGACCTAA AGGTTATGAC ATTTTTACTG
                  Consensus  ********** ********** ********** ********** **********

                             1601                                              1650
      msa178297.2{104_090}   TTGGGATTGG TGTAAACGGA GATCCTGGTG CAACTCCTCT TGAAGCTGAG
    msa178297.2{104_18RS21}  TTGGGATTGG TGTAAACGGA GATCCTGGTG CAACTCCTCT TGAAGCTGAG
     msa178297.2{104_2603}   TTGGGATTGG TGTAAACGGA GATCCTGGTG CAACTCCTCT TGAAGCTGAG
    msa178297.2{104_CJB110}  TTGGGATTGG TGTAAACGGA GATCCTGGTG CAACTCCTCT TGAAGCTGAG
     msa178297.2{104_COH1}   TTGGGATTGG TGTAAACGGA GATCCTGGTG CAACTCCTCT TGAAGCTGAG
     msa178297.2{104_M732}   TTGGGATTGG TGTAAACGGA GATCCTGGTG CAACTCCTCT TGAAGCTGAG
     msa178297.2{104_A909}   TTGGGATTGG TGTAAACGGA GATCCTGGTG CAACTCCTCT TGAAGCTGAG
     msa178297.2{104_M781}   TTGGGATTGG TGTAAACGGA GATCCTGGTG CAACTCCTCT TGAAGCTGAG
 msa178297.2{104_JM9130013}  TTGGGATTGG TGTAAACGGA GATCCTGGTG CAACTCCTCT TGAAGCTGAG
                  Consensus  ********** ********** ********** ********** **********

                             1651                                              1700
      msa178297.2{104_090}   AAATTTATGC AATCAATATC AAGTAAAACA GAAAATTATA CTAATGTTGA
    msa178297.2{104_18RS21}  AAATTTATGC AATCAATATC AAGTAAAACA GAAAATTATA CTAATGTTGA
     msa178297.2{104_2603}   AAATTTATGC AATCAATATC AAGTAAAACA GAAAATTATA CTAATGTTGA
    msa178297.2{104_CJB110}  AAATTTATGC AATCAATATC AAGTAAAACA GAAAATTATA CTAATGTTGA
     msa178297.2{104_COH1}   AAATTTATGC AATCAATATC AAGTAAAACA GAAAATTATA CTAATGTTGA
     msa178297.2{104_M732}   AAATTTATGC AATCAATATC AAGTAAAACA GAAAATTATA CTAATGTTGA
     msa178297.2{104_A909}   AAATTTATGC AATCAATATC AAGTAAAACA GAAAATTATA CTAATGTTGA
     msa178297.2{104_M781}   AAATTTATGC AATCAATATC AAGTAAAACA GAAAATTATA CTAATGTTGA
 msa178297.2{104_JM9130013}  AAATTTATGC AATCAATATC AAGTAAAACA GAAAATTATA CTAATGTTGA
                  Consensus  ********** ********** ********** ********** **********

                             1701                                              1750
      msa178297.2{104_090}   TGATACAAAT AAAATTTATG ATGAGCTAAA TAAATACTTT AAAACAATTG
    msa178297.2{104_18RS21}  TGATACAAAT AAAATTTATG ATGAGCTAAA TAAATACTTT AAAACAATTG
     msa178297.2{104_2603}   TGATACAAAT AAAATTTATG ATGAGCTAAA TAAATACTTT AAAACAATTG
    msa178297.2{104_CJB110}  TGATACAAAT AAAATTTATG ATGAGCTAAA TAAATACTTT AAAACAATTG
```

Table 41: Comparative Sequences relating to SAG0649

```
       msa178297.2{104_COH1}  TGATACAAAT AAAATTTATG ATGAGCTAAA TAAATACTTT AAAACAATTG
       msa178297.2{104_M732}  TGATACAAAT AAAATTTATG ATGAGCTAAA TAAATACTTT AAAACAATTG
       msa178297.2{104_A909}  TGATACAAAT AAAATTTATG ATGAGCTAAA TAAATACTTT AAAACAATTG
       msa178297.2{104_M781}  TGATACAAAT AAAATTTATG ATGAGCTAAA TAAATACTTT AAAACAATTG
  msa178297.2{104_JM9130013}  TGATACAAAT AAAATTTATG ATGAGCTAAA TAAATACTTT AAAACAATTG
                   Consensus  ********** ********** ********** ********** **********

                              1751                                               1800
        msa178297.2{104_090}  TTGAGGAAAA ACATTCTATT GTTGATGGAA ATGTGACTGA TCCTATGGGA
     msa178297.2{104_18RS21}  TTGAGGAAAA ACATTCTATT GTTGATGGAA ATGTGACTGA TCCTATGGGA
       msa178297.2{104_2603}  TTGAGGAAAA ACATTCTATT GTTGATGGAA ATGTGACTGA TCCTATGGGA
     msa178297.2{104_CJB110}  TTGAGGAAAA ACATTCTATT GTTGATGGAA ATGTGACTGA TCCTATGGGA
       msa178297.2{104_COH1}  TTGAGGAAAA ACATTCTATT GTTGATGGAA ATGTGACTGA TCCTATGGGA
       msa178297.2{104_M732}  TTGAGGAAAA ACATTCTATT GTTGATGGAA ATGTGACTGA TCCTATGGGA
       msa178297.2{104_A909}  TTGAGGAAAA ACATTCTATT GTTGATGGAA ATGTGACTGA TCCTATGGGA
       msa178297.2{104_M781}  TTGAGGAAAA ACATTCTATT GTTGATGGAA ATGTGACTGA TCCTATGGGA
  msa178297.2{104_JM9130013}  TTGAGGAAAA ACATTCTATT GTTGATGGAA ATGTGACTGA TCCTATGGGA
                   Consensus  ********** ********** ********** ********** **********

                              1801                                               1850
        msa178297.2{104_090}  GAGATGATTG AATTCCAATT AAAAAATGGT CAAAGTTTTA CACATGATGA
     msa178297.2{104_18RS21}  GAGATGATTG AATTCCAATT AAAAAATGGT CAAAGTTTTA CACATGATGA
       msa178297.2{104_2603}  GAGATGATTG AATTCCAATT AAAAAATGGT CAAAGTTTTA CACATGATGA
     msa178297.2{104_CJB110}  GAGATGATTG AATTCCAATT AAAAAATGGT CAAAGTTTTA CACATGATGA
       msa178297.2{104_COH1}  GAGATGATTG AATTCCAATT AAAAAATGGT CAAAGTTTTA CACATGATGA
       msa178297.2{104_M732}  GAGATGATTG AATTCCAATT AAAAAATGGT CAAAGTTTTA CACATGATGA
       msa178297.2{104_A909}  GAGATGATTG AATTCCAATT AAAAAATGGT CAAAGTTTTA CACATGATGA
       msa178297.2{104_M781}  GAGATGATTG AATTCCAATT AAAAAATGGT CAAAGTTTTA CACATGATGA
  msa178297.2{104_JM9130013}  GAGATGATTG AATTCCAATT AAAAAATGGT CAAAGTTTTA CACATGATGA
                   Consensus  ********** ********** ********** ********** **********

                              1851                                               1900
        msa178297.2{104_090}  TTACGTTTTG GTTGGAAATG ATGGCAGTCA ATTAAAAAAT GGTGTGGCTC
     msa178297.2{104_18RS21}  TTACGTTTTG GTTGGAAATG ATGGCAGTCA ATTAAAAAAT GGTGTGGCTC
       msa178297.2{104_2603}  TTACGTTTTG GTTGGAAATG ATGGCAGTCA ATTAAAAAAT GGTGTGGCTC
     msa178297.2{104_CJB110}  TTACGTTTTG GTTGGAAATG ATGGCAGTCA ATTAAAAAAT GGTGTGGCTC
       msa178297.2{104_COH1}  TTACGTTTTG GTTGGAAATG ATGGCAGTCA ATTAAAAAAT GGTGTGGCTC
       msa178297.2{104_M732}  TTACGTTTTG GTTGGAAATG ATGGCAGTCA ATTAAAAAAT GGTGTGGCTC
       msa178297.2{104_A909}  TTACGTTTTG GTTGGAAATG ATGGCAGTCA ATTAAAAAAT GGTGTGGCTC
       msa178297.2{104_M781}  TTACGTTTTG GTTGGAAATG ATGGCAGTCA ATTAAAAAAT GGTGTGGCTC
  msa178297.2{104_JM9130013}  TTACGTTTTG GTTGGAAATG ATGGCAGTCA ATTAAAAAAT GGTGTGGCTC
                   Consensus  ********** ********** ********** ********** **********

                              1901                                               1950
        msa178297.2{104_090}  TTGGTGGACC AAACAGTGAT GGGGGAATTT TAAAAGATGT TACAGTGACT
     msa178297.2{104_18RS21}  TTGGTGGACC AAACAGTGAT GGGGGAATTT TAAAAGATGT TACAGTGACT
       msa178297.2{104_2603}  TTGGTGGACC AAACAGTGAT GGGGGAATTT TAAAAGATGT TACAGTGACT
     msa178297.2{104_CJB110}  TTGGTGGACC AAACAGTGAT GGGGGAATTT TAAAAGATGT TACAGTGACT
       msa178297.2{104_COH1}  TTGGTGGACC AAACAGTGAT GGGGGAATTT TAAAAGATGT TACAGTGACT
       msa178297.2{104_M732}  TTGGTGGACC AAACAGTGAT GGGGGAATTT TAAAAGATGT TACAGTGACT
       msa178297.2{104_A909}  TTGGTGGACC AAACAGTGAT GGGGGAATTT TAAAAGATGT TACAGTGACT
       msa178297.2{104_M781}  TTGGTGGACC AAACAGTGAT GGGGGAATTT TAAAAGATGT TACAGTGACT
  msa178297.2{104_JM9130013}  TTGGTGGACC AAACAGTGAT GGGGGAATTT TAAAAGATGT TACAGTGACT
                   Consensus  ********** ********** ********** ********** **********

                              1951                                               2000
        msa178297.2{104_090}  TATGATAAGA CATCTCAAAC CATCAAAATC AATCATTTGA ACTTAGGAAG
     msa178297.2{104_18RS21}  TATGATAAGA CATCTCAAAC CATCAAAATC AATCATTTGA ACTTAGGAAG
       msa178297.2{104_2603}  TATGATAAGA CATCTCAAAC CATCAAAATC AATCATTTGA ACTTAGGAAG
     msa178297.2{104_CJB110}  TATGATAAGA CATCTCAAAC CATCAAAATC AATCATTTGA ACTTAGGAAG
       msa178297.2{104_COH1}  TATGATAAGA CATCTCAAAC CATCAAAATC AATCATTTGA ACTTAGGAAG
       msa178297.2{104_M732}  TATGATAAGA CATCTCAAAC CATCAAAATC AATCATTTGA ACTTAGGAAG
       msa178297.2{104_A909}  TATGATAAGA CATCTCAAAC CATCAAAATC AATCATTTGA ACTTAGGAAG
       msa178297.2{104_M781}  TATGATAAGA CATCTCAAAC CATCAAAATC AATCATTTGA ACTTAGGAAG
  msa178297.2{104_JM9130013}  TATGATAAGA CATCTCAAAC CATCAAAATC AATCATTTGA ACTTAGGAAG
                   Consensus  ********** ********** ********** ********** **********

                              2001                                               2050
        msa178297.2{104_090}  TGGACAAAAA GTAGTTCTTA CCTATGATGT ACGTTTAAAA GATAACTATA
     msa178297.2{104_18RS21}  TGGACAAAAA GTAGTTCTTA CCTATGATGT ACGTTTAAAA GATAACTATA
       msa178297.2{104_2603}  TGGACAAAAA GTAGTTCTTA CCTATGATGT ACGTTTAAAA GATAACTATA
     msa178297.2{104_CJB110}  TGGACAAAAA GTAGTTCTTA CCTATGATGT ACGTTTAAAA GATAACTATA
       msa178297.2{104_COH1}  TGGACAAAAA GTAGTTCTTA CCTATGATGT ACGTTTAAAA GATAACTATA
       msa178297.2{104_M732}  TGGACAAAAA GTAGTTCTTA CCTATGATGT ACGTTTAAAA GATAACTATA
       msa178297.2{104_A909}  TGGACAAAAA GTAGTTCTTA CCTATGATGT ACGTTTAAAA GATAACTATA
       msa178297.2{104_M781}  TGGACAAAAA GTAGTTCTTA CCTATGATGT ACGTTTAAAA GATAACTATA
  msa178297.2{104_JM9130013}  TGGACAAAAA GTAGTTCTTA CCTATGATGT ACGTTTAAAA GATAACTATA
                   Consensus  ********** ********** ********** ********** **********

                              2051                                               2100
        msa178297.2{104_090}  TAAGTAACAA ATTTTACAAT ACAAATAATC GTACAACGCT AAGTCCGAAG
     msa178297.2{104_18RS21}  TAAGTAACAA ATTTTACAAT ACAAATAATC GTACAACGCT AAGTCCGAAG
       msa178297.2{104_2603}  TAAGTAACAA ATTTTACAAT ACAAATAATC GTACAACGCT AAGTCCGAAG
```

Table 41: Comparative Sequences relating to SAG0649

```
     msa178297.2{104_CJB110}  TAAGTAACAA ATTTTACAAT ACAAATAATC GTACAACGCT AAGTCCGAAG
       msa178297.2{104_COH1}  TAAGTAACAA ATTTTACAAT ACAAATAATC GTACAACGCT AAGTCCGAAG
       msa178297.2{104_M732}  TAAGTAACAA ATTTTACAAT ACAAATAATC GTACAACGCT AAGTCCGAAG
       msa178297.2{104_A909}  TAAGTAACAA ATTTTACAAT ACAAATAATC GTACAACGCT AAGTCCGAAG
       msa178297.2{104_M781}  TAAGTAACAA ATTTTACAAT ACAAATAATC GTACAACGCT AAGTCCGAAG
 msa178297.2{104_JM9130013}  TAAGTAACAA ATTTTACAAT ACAAATAATC GTACAACGCT AAGTCCGAAG
                   Consensus  ********** ********** ********** ********** **********

                              2101                                               2150
        msa178297.2{104_090}  AGTGAAAAAG AACCAAATAC TATTCGTGAT TTCCCAATTC CCAAAATTCG
     msa178297.2{104_18RS21}  AGTGAAAAAG AACCAAATAC TATTCGTGAT TTCCCAATTC CCAAAATTCG
       msa178297.2{104_2603}  AGTGAAAAAG AACCAAATAC TATTCGTGAT TTCCCAATTC CCAAAATTCG
     msa178297.2{104_CJB110}  AGTGAAAAAG AACCAAATAC TATTCGTGAT TTCCCAATTC CCAAAATTCG
       msa178297.2{104_COH1}  AGTGAAAAAG AACCAAATAC TATTCGTGAT TTCCCAATTC CCAAAATTCG
       msa178297.2{104_M732}  AGTGAAAAAG AACCAAATAC TATTCGTGAT TTCCCAATTC CCAAAATTCG
       msa178297.2{104_A909}  AGTGAAAAAG AACCAAATAC TATTCGTGAT TTCCCAATTC CCAAAATTCG
       msa178297.2{104_M781}  AGTGAAAAAG AACCAAATAC TATTCGTGAT TTCCCAATTC CCAAAATTCG
  msa178297.2{104_JM9130013}  AGTGAAAAAG AACCAAATAC TATTCGTGAT TTCCCAATTC CCAAAATTCG
                   Consensus  ********** ********** ********** ********** **********

                              2151                                               2200
        msa178297.2{104_090}  TGATGTTCGT GAGTTTCCGG TACTAACCAT CAGTAATCAg AAGAAAATGG
     msa178297.2{104_18RS21}  TGATGTTCGT GAGTTTCCGG TACTAACCAT CAGTAATCAg AAGAAAATGG
       msa178297.2{104_2603}  TGATGTTCGT GAGTTTCCGG TACTAACCAT CAGTAATCAg AAGAAAATGG
     msa178297.2{104_CJB110}  TGATGTTCGT GAGTTTCCGG TACTAACCAT CAGTAATCAg AAGAAAATGG
       msa178297.2{104_COH1}  TGATGTTCGT GAGTTTCCGG TACTAACCAT CAGTAATCAg AAGAAAATGG
       msa178297.2{104_M732}  TGATGTTCGT GAGTTTCCGG TACTAACCAT CAGTAATCAg AAGAAAATGG
       msa178297.2{104_A909}  TGATGTTCGT GAGTTTCCGG TACTAACCAT CAGTAATCAg AAGAAAATGG
       msa178297.2{104_M781}  TGATGTTCGT GAGTTTCCGG TACTAACCAT CAGTAATCAg AAGAAAATGG
  msa178297.2{104_JM9130013}  TGATGTTCGT GAGTTTCCGG TACTAACCAT CAGTAATCAa AAGAAAATGG
                   Consensus  ********** ********** ********** *********_ **********

                              2201                                               2250
        msa178297.2{104_090}  GTGAGGTTGA ATTTATTAAA GTTAATAAAG ACAAACATTC AGAATCGCTT
     msa178297.2{104_18RS21}  GTGAGGTTGA ATTTATTAAA GTTAATAAAG ACAAACATTC AGAATCGCTT
       msa178297.2{104_2603}  GTGAGGTTGA ATTTATTAAA GTTAATAAAG ACAAACATTC AGAATCGCTT
     msa178297.2{104_CJB110}  GTGAGGTTGA ATTTATTAAA GTTAATAAAG ACAAACATTC AGAATCGCTT
       msa178297.2{104_COH1}  GTGAGGTTGA ATTTATTAAA GTTAATAAAG ACAAACATTC AGAATCGCTT
       msa178297.2{104_M732}  GTGAGGTTGA ATTTATTAAA GTTAATAAAG ACAAACATTC AGAATCGCTT
       msa178297.2{104_A909}  GTGAGGTTGA ATTTATTAAA GTTAATAAAG ACAAACATTC AGAATCGCTT
       msa178297.2{104_M781}  GTGAGGTTGA ATTTATTAAA GTTAATAAAG ACAAACATTC AGAATCGCTT
  msa178297.2{104_JM9130013}  GTGAGGTTGA ATTTATTAAA GTTAATAAAG ACAAACATTC AGAATCGCTT
                   Consensus  ********** ********** ********** ********** **********

                              2251                                               2300
        msa178297.2{104_090}  TTGGGAGCTA AGTTTCAACT TCAGATAgAA AAAGATTTTT CTGGGTATAA
     msa178297.2{104_18RS21}  TTGGGAGCTA AGTTTCAACT TCAGATAgAA AAAGATTTTT CTGGGTATAA
       msa178297.2{104_2603}  TTGGGAGCTA AGTTTCAACT TCAGATAgAA AAAGATTTTT CTGGGTATAA
     msa178297.2{104_CJB110}  TTGGGAGCTA AGTTTCAACT TCAGATAgAA AAAGATTTTT CTGGGTATAA
       msa178297.2{104_COH1}  TTGGGAGCTA AGTTTCAACT TCAGATAgAA AAAGATTTTT CTGGGTATAA
       msa178297.2{104_M732}  TTGGGAGCTA AGTTTCAACT TCAGATAgAA AAAGATTTTT CTGGGTATAA
       msa178297.2{104_A909}  TTGGGAGCTA AGTTTCAACT TCAGATAgAA AAAGATTTTT CTGGGTATAA
       msa178297.2{104_M781}  TTGGGAGCTA AGTTTCAACT TCAGATAgAA AAAGATTTTT CTGGGTATAA
  msa178297.2{104_JM9130013}  TTGGGAGCTA AGTTTCAACT TCAGATAaAA AAAGATTTTT CTGGGTATAA
                   Consensus  ********** ********** *******_** ********** **********

                              2301                                               2350
        msa178297.2{104_090}  GCAATTTGTT CCAGAGGGAA GTGATGTTAC AACAAAGAAT GATGGTAAAA
     msa178297.2{104_18RS21}  GCAATTTGTT CCAGAGGGAA GTGATGTTAC AACAAAGAAT GATGGTAAAA
       msa178297.2{104_2603}  GCAATTTGTT CCAGAGGGAA GTGATGTTAC AACAAAGAAT GATGGTAAAA
     msa178297.2{104_CJB110}  GCAATTTGTT CCAGAGGGAA GTGATGTTAC AACAAAGAAT GATGGTAAAA
       msa178297.2{104_COH1}  GCAATTTGTT CCAGAGGGAA GTGATGTTAC AACAAAGAAT GATGGTAAAA
       msa178297.2{104_M732}  GCAATTTGTT CCAGAGGGAA GTGATGTTAC AACAAAGAAT GATGGTAAAA
       msa178297.2{104_A909}  GCAATTTGTT CCAGAGGGAA GTGATGTTAC AACAAAGAAT GATGGTAAAA
       msa178297.2{104_M781}  GCAATTTGTT CCAGAGGGAA GTGATGTTAC AACAAAGAAT GATGGTAAAA
  msa178297.2{104_JM9130013}  GCAATTTGTT CCAGAGGGAA GTGATGTTAC AACAAAGAAT GATGGTAAAA
                   Consensus  ********** ********** ********** ********** **********

                              2351                                               2400
        msa178297.2{104_090}  TTTATTTTAA AGCACTTCAA GATGGTAACT ATAAATTATA TGAAATTTCA
     msa178297.2{104_18RS21}  TTTATTTTAA AGCACTTCAA GATGGTAACT ATAAATTATA TGAAATTTCA
       msa178297.2{104_2603}  TTTATTTTAA AGCACTTCAA GATGGTAACT ATAAATTATA TGAAATTTCA
     msa178297.2{104_CJB110}  TTTATTTTAA AGCACTTCAA GATGGTAACT ATAAATTATA TGAAATTTCA
       msa178297.2{104_COH1}  TTTATTTTAA AGCACTTCAA GATGGTAACT ATAAATTATA TGAAATTTCA
       msa178297.2{104_M732}  TTTATTTTAA AGCACTTCAA GATGGTAACT ATAAATTATA TGAAATTTCA
       msa178297.2{104_A909}  TTTATTTTAA AGCACTTCAA GATGGTAACT ATAAATTATA TGAAATTTCA
       msa178297.2{104_M781}  TTTATTTTAA AGCACTTCAA GATGGTAACT ATAAATTATA TGAAATTTCA
  msa178297.2{104_JM9130013}  TTTATTTTAA AGCACTTCAA GATGGTAACT ATAAATTATA TGAAATTTCA
                   Consensus  ********** ********** ********** ********** **********

                              2401                                               2450
        msa178297.2{104_090}  AGTCCAGATG GCTATATAGA GGTTAAAACG AAACCTGTTG TGACATTTAC
     msa178297.2{104_18RS21}  AGTCCAGATG GCTATATAGA GGTTAAAACG AAACCTGTTG TGACATTTAC
```

Table 41: Comparative Sequences relating to SAG0649

```
       msa178297.2{104_2603}  AGTCCAGATG GCTATATAGA GGTTAAAACG AAACCTGTTG TGACATTTAC
     msa178297.2{104_CJB110}  AGTCCAGATG GCTATATAGA GGTTAAAACG AAACCTGTTG TGACATTTAC
       msa178297.2{104_COH1}  AGTCCAGATG GCTATATAGA GGTTAAAACG AAACCTGTTG TGACATTTAC
       msa178297.2{104_M732}  AGTCCAGATG GCTATATAGA GGTTAAAACG AAACCTGTTG TGACATTTAC
       msa178297.2{104_A909}  AGTCCAGATG GCTATATAGA GGTTAAAACG AAACCTGTTG TGACATTTAC
       msa178297.2{104_M781}  AGTCCAGATG GCTATATAGA GGTTAAAACG AAACCTGTTG TGACATTTAC
  msa178297.2{104_JM9130013}  AGTCCAGATG GCTATATAGA GGTTAAAACG AAACCTGTTG TGACATTTAC
                   Consensus  ********** ********** ********** ********** **********

                              2451                                           2500
        msa178297.2{104_090}  AATTCAAAAT GGAGAAGTTA CGAACCTGAA AGCAGATCCA AATGCTAATA
      msa178297.2{104_18RS21}  AATTCAAAAT GGAGAAGTTA CGAACCTGAA AGCAGATCCA AATGCTAATA
       msa178297.2{104_2603}  AATTCAAAAT GGAGAAGTTA CGAACCTGAA AGCAGATCCA AATGCTAATA
     msa178297.2{104_CJB110}  AATTCAAAAT GGAGAAGTTA CGAACCTGAA AGCAGATCCA AATGCTAATA
       msa178297.2{104_COH1}  AATTCAAAAT GGAGAAGTTA CGAACCTGAA AGCAGATCCA AATGCTAATA
       msa178297.2{104_M732}  AATTCAAAAT GGAGAAGTTA CGAACCTGAA AGCAGATCCA AATGCTAATA
       msa178297.2{104_A909}  AATTCAAAAT GGAGAAGTTA CGAACCTGAA AGCAGATCCA AATGCTAATA
       msa178297.2{104_M781}  AATTCAAAAT GGAGAAGTTA CGAACCTGAA AGCAGATCCA AATGCTAATA
  msa178297.2{104_JM9130013}  AATTCAAAAT GGAGAAGTTA CGAACCTGAA AGCAGATCCA AATGCTAATA
                   Consensus  ********** ********** ********** ********** **********

                              2501                                           2550
        msa178297.2{104_090}  AAAATCAAAT CGGGTATCTT GAAggaaatg gtaaacatct tattaccaac
      msa178297.2{104_18RS21}  AAAATCAAAT CGGGTATCTT GAAggaaatg gtaaacatct tattaccaac
       msa178297.2{104_2603}  AAAATCAAAT CGGGTATCTT GAAggaaatg gtaaacatct tattaccaac
     msa178297.2{104_CJB110}  AAAATCAAAT CGGGTATCTT GAAggaaatg gtaaacatct tattaccaac
       msa178297.2{104_COH1}  AAAATCAAAT CGGGTATCTT GAAggaaatg gtaaacatct tattaccaac
       msa178297.2{104_M732}  AAAATCAAAT CGGGTATCTT GAAggaaatg gtaaacatct tattaccaac
       msa178297.2{104_A909}  AAAATCAAAT CGGGTATCTT GAAggaaatg gtaaacatct tattaccaac
       msa178297.2{104_M781}  AAAATCAAAT CGGGTATCTT GAAggaaatg gtaaacatct tattaccaac
  msa178297.2{104_JM9130013}  AAAATCAAAT CGGGTATCTT GAA~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
                   Consensus  ********** ********** ***------- ---------- ----------

                              2551                                           2600
        msa178297.2{104_090}  actcccaaac gcccaccagg tgtt~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      msa178297.2{104_18RS21}  actcccaaac gcccaccagg tgtt~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_2603}  actcccaaac gcccaccagg tgtttttcct aaaacagggg gaattggtac
     msa178297.2{104_CJB110}  actcccaaac gcccaccagg tgtt~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_COH1}  actcccaaac gcccaccagg tgtt~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_M732}  actcccaaac gcccaccagg tgtt~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_A909}  actcccaaac gcccaccagg tgtt~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_M781}  actcccaaac gcccaccagg tgtt~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  msa178297.2{104_JM9130013}  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
                   Consensus  ---------- ---------- ----****** ********** **********

                              2601                                           2650
        msa178297.2{104_090}  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      msa178297.2{104_18RS21}  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_2603}  aattgtctat atattagttg gttctacttt tatgatactt accatttgtt
     msa178297.2{104_CJB110}  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_COH1}  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_M732}  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_A909}  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_M781}  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  msa178297.2{104_JM9130013}  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
                   Consensus  ********** ********** ********** ********** **********

                              2651              2670
        msa178297.2{104_090}  ~~~~~~~~~~ ~~~~~~~~~~
      msa178297.2{104_18RS21}  ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_2603}  ctttccgtcg taaacaattg
     msa178297.2{104_CJB110}  ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_COH1}  ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_M732}  ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_A909}  ~~~~~~~~~~ ~~~~~~~~~~
       msa178297.2{104_M781}  ~~~~~~~~~~ ~~~~~~~~~~
  msa178297.2{104_JM9130013}  ~~~~~~~~~~ ~~~~~~~~~~
                   Consensus  ********** **********
```

Table 41: Comparative Sequences relating to SAG0649

SEQ ID NO. 4110: SAG0649 FROM 2603 V/R GBS TYPE V STRAIN
MKKRQKIWRGLSVTLLILSQIPFGILVQGETQDTNQALGKVIVKKTGDNATPLGKATFVL
KNDNDKSETSHETVEGSGEATFENIKPGDYTLREETAPIGYKKTDKTWKVKVADNGATII
EGMDADKAEKRKEVLNAQYPKSAIYEDTKENYPLVNVEGSKVGEQYKALNPINGKDGRRE
IAEGWLSKKITGVNDLDKNKYKIELTVEGKTTVETKELNQPLDVVVLLDNSNSMNNERAN
NSQRALKAGEAVEKLIDKITSNKDNRVALVTYASTIFDGTEATVSKGVADQNGKALNDSV
SWDYHKTTFTATTHNYSYLNLTNDANEVNILKSRIPKEAEHINGDRTLYQFGATFTQKAL
MKANEILETQSSNARKKLIFHVTDGVPTMSYAINFNPYISTSYQNQFNSFLNKIPDRSGI
LQEDFIINGDDYQIVKGDGESFKLFSDRKVPVTGGTTQAAYRVPQNQLSVMSNEGYAINS
GYIYLYWRDYNWVYPFDPKTKKVSATKQIKTHGEPTTLYFNGNIRPKGYDIFTVGIGVNG
DPGATPLEAEKFMQSISSKTENYTNVDDTNKIYDELNKYFKTIVEEKHSIVDGNVTDPMG
EMIEFQLKNGQSFTHDDYVLVGNDGSQLKNGVALGGPNSDGGILKDVTVTYDKTSQTIKI
NHLNLGSGQKVVLTYDVRLKDNYISNKFYNTNNRTTLSPKSEKEPNTIRDFPIPKIRDVR
EFPVLTISNQKKMGEVEFIKVNKDKHSESLLGAKFQLQIEKDFSGYKQFVPEGSDVTTKN
DGKIYFKALQDGNYKLYEISSPDGYIEVKTKPVVTFTIQNGEVTNLKADPNANKNQIGYL
EGNGKHLITNTPKRPPGVFPKTGGIGTIVYILVGSTFMILTICSFRRKQL

SEQ ID NO. 4111: SAG0649 FROM 090 GBS TYPE Ia STRAIN
GETQDTNQALGKVIVKKTGDNATPLGKATFVLKNDNDKSETSHETVEGSGEATFENIKPG
DYTLREETAPIGYKKTDKTWKVKVADNGATIIEGMDADKAEKRKEVLNAQYPKSAIYEDT
KENYPLVNVEGSKVGEQYKALNPINGKDGRREIAEGWLSKKITGVNDLDKNKYKIELTVE
GKTTVETKELNQPLDVVVLLDNSNSMNNERANNSQRALKAGEAVEKLIDKITSNKDNRVA
LVTYASTIFDGTEATVSKGVADQNGKALNDSVSWDYHKTTFTATTHNYSYLNLTNDANEV
NILKSRIPKEAEHINGDRTLYQFGATFTQKALMKANEILETQSSNARKKLIFHVTDGVPT
MSYAINFNPYISTSYQNQFNSFLNKIPDRSGILQEDFIINGDDYQIVKGDGESFKLFSDR
KVPVTGGTTQAAYRVPQNQLSVMSNEGYAINSGYIYLYWRDYNWVYPFDPKTKKVSATKQ
IKTHGEPTTLYFNGNIRPKGYDIFTVGIGVNGDPGATPLEAEKFMQSISSKTENYTNVDD
TNKIYDELNKYFKTIVEEKHSIVDGNVTDPMGEMIEFQLKNGQSFTHDDYVLVGNDGSQL
KNGVALGGPNSDGGILKDVTVTYDKTSQTIKINHLNLGSGQKVVLTYDVRLKDNYISNKF
YNTNNRTTLSPKSEKEPNTIRDFPIPKIRDVREFPVLTISNQKKMGEVEFIKVNKDKHSE
SLLGAKFQLQIEKDFSGYKQFVPEGSDVTTKNDGKIYFKALQDGNYKLYEISSPDGYIEV
KTKPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITNTPKRPPGV

SEQ ID NO. 4112: SAG0649 FROM A909 GBS TYPE Ia STRAIN
GETQDTNQALGKVIVKKTGDNATPLGKATFVLKNDNDKSETSHETVEGSGEATFENIKPG
DYTLREETAPIGYKKTDKTWKVKVADNGATIIEGMDADKAEKRKEVLNAQYPKSAIYEDT
KENYPLVNVEGSKVGEQYKALNPINGKDGRREIAEGWLSKKITGVNDLDKNKYKIELTVE
GKTTVETKELNQPLDVVVLLDNSNSMNNERANNSQRALKAGEAVEKLIDKITSNKDNRVA
LVTYASTIFDGTEATVSKGVADQNGKALNDSVSWDYHKTTFTATTHNYSYLNLTNDANEV
NILKSRIPKEAEHINGDRTLYQFGATFTQKALMKANEILETQSSNARKKLIFHVTDGVPT
MSYAINFNPYISTSYQNQFNSFLNKIPDRSGILQEDFIINGDDYQIVKGDGESFKLFSDR
KVPVTGGTTQAAYRVPQNQLSVMSNEGYAINSGYIYLYWRDYNWVYPFDPKTKKVSATKQ
IKTHGEPTTLYFNGNIRPKGYDIFTVGIGVNGDPGATPLEAEKFMQSISSKTENYTNVDD
TNKIYDELNKYFKTIVEEKHSIVDGNVTDPMGEMIEFQLKNGQSFTHDDYVLVGNDGSQL
KNGVALGGPNSDGGILKDVTVTYDKTSQTIKINHLNLGSGQKVVLTYDVRLKDNYISNKF
YNTNNRTTLSPKSEKEPNTIRDFPIPKIRDVREFPVLTISNQKKMGEVEFIKVNKDKHSE
SLLGAKFQLQIEKDFSGYKQFVPEGSDVTTKNDGKIYFKALQDGNYKLYEISSPDGYIEV
KTKPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITNTPKRPPGV

SEQ ID NO. 4113: SAG0649 FROM 18RS21 GBS TYPE II STRAIN
GETQDTNQALGKVIVKKTGDNATPLGKATFVLKNDNDKSETSHETVEGSGEATFENIKPG
DYTLREETAPIGYKKTDKTWKVKVADNGATIIEGMDADKAEKRKEVLNAQYPKSAIYEDT
KENYPLVNVEGSKVGEQYKALNPINGKDGRREIAEGWLSKKITGVNDLDKNKYKIELTVE
GKTTVETKELNQPLDVVVLLDNSNSMNNERANNSQRALKAGEAVEKLIDKITSNKDNRVA
LVTYASTIFDGTEATVSKGVADQNGKALNDSVSWDYHKTTFTATTHNYSYLNLTNDANEV
NILKSRIPKEAEHINGDRTLYQFGATFTQKALMKANEILETQSSNARKKLIFHVTDGVPT
MSYAINFNPYISTSYQNQFNSFLNKIPDRSGILQEDFIINGDDYQIVKGDGESFKLFSDR
KVPVTGGTTQAAYRVPQNQLSVMSNEGYAINSGYIYLYWRDYNWVYPFDPKTKKVSATKQ
IKTHGEPTTLYFNGNIRPKGYDIFTVGIGVNGDPGATPLEAEKFMQSISSKTENYTNVDD
TNKIYDELNKYFKTIVEEKHSIVDGNVTDPMGEMIEFQLKNGQSFTHDDYVLVGNDGSQL
KNGVALGGPNSDGGILKDVTVTYDKTSQTIKINHLNLGSGQKVVLTYDVRLKDNYISNKF
YNTNNRTTLSPKSEKEPNTIRDFPIPKIRDVREFPVLTISNQKKMGEVEFIKVNKDKHSE
SLLGAKFQLQIEKDFSGYKQFVPEGSDVTTKNDGKIYFKALQDGNYKLYEISSPDGYIEV
KTKPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITNTPKRPPGV

SEQ ID NO. 4114: SAG0649 FROM M732 GBS TYPE III STRAIN
GETQDTNQALGKVIVKKTGDNATPLGKATFVLKNDNDKSETSHETVEGSGEATFENIKPG
DYTLREETAPIGYKKTDKTWKVKVADNGATIIEGMDADKAEKRKEVLNAQYPKSAIYEDT
KENYPLVNVEGSKVGEQYKALNPINGKDGRREIAEGWLSKKNTGVNDLDKNKYKIELTVE
GKTTVETKELNQPLDVVVLLDNSNSMNNERANNSQRALKAGEAVEKLIDKITSNKDNRVA
LVTYASTIFDGTEATVSKGVADQNGKALNDSVSWDYHKTTFTATTHNYSYLNLTNDANEV
NILKSRIPKEAEHINGDRTLYQFGATFTQKALMKANEILETQSSNARKKLIFHVTDGVPT
MSYAINFNPYISTSYQNQFNSFLNKIPDRSGILQEDFIINGDDYQIVKGDGESFKLFSDR
KVPVTGGTTQAAYRVPQNQLSVMSNEGYAINSGYIYLYWRDYNWVYPFDPKTKKVSATKQ
IKTHGEPTTLYFNGNIRPKGYDIFTVGIGVNGDPGATPLEAEKFMQSISSKTENYTNVDD
TNKIYDELNKYFKTIVEEKHSIVDGNVTDPMGEMIEFQLKNGQSFTHDDYVLVGNDGSQL
KNGVALGGPNSDGGILKDVTVTYDKTSQTIKINHLNLGSGQKVVLTYDVRLKDNYISNKF
YNTNNRTTLSPKSEKEPNTIRDFPIPKIRDVREFPVLTISNQKKMGEVEFIKVNKDKHSE
SLLGAKFQLQIEKDFSGYKQFVPEGSDVTTKNDGKIYFKALQDGNYKLYEISSPDGYIEV
KTKPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITNTPKRPPGV

SEQ ID NO. 4115: SAG0649 FROM COH1 GBS TYPE III STRAIN
GETQDTNQALGKVIVKKTGDNATPLGKATFVLKNDNDKSETSHETVEGSGXATFENIKPG

Table 41: Comparative Sequences relating to SAG0649

DYTLREETAPIGYKKTDKTWKVKVADNGATIIEGMDADKAEKRKEVLNAQYPKSAIYEDT
KENYPLVNVEGSKVGEQYKALNPINGKDGRREIAEGWLSKKNTGVNDLDKNKYKIELTVE
GKTTVETKELNQPLDVVVLLDNSNSMNNERANNSQRALKAGEAVEKLIDKITSNKDNRVA
LVTYASTIFDGTEATVSKGVADQNGKALNDSVSWDYHKTTFTATTHNYSYLNLTNDANEV
NILKSRIPKEAEHINGDRTLYQFGATFTQKALMKANEILETQSSNARKKLIFHVTDGVPT
MSYAINFNPYISTSYQNQFNSFLNKIPDRSGILQEDFIINGDDYQIVKGDGESFKLFSDR
KVPVTGGTTQAAYRVPQNQLSVMSNEGYAINSGYIYLYWRDYNWVYPFDPKTKKVSATKQ
IKTHGEPTTLYFNGNIRPKGYDIFTVGIGVNGDPGATPLEAEKFMQSISSKTENYTNVDD
TNKIYDELNKYFKTIVEEKHSIVDGNVTDPMGEMIEFQLKNGQSFTHDDYVLVGNDGSQL
KNGVALGGPNSDGGILKDVTVTYDKTSQTIKINHLNLGSGQKVVLTYDVRLKDNYISNKF
YNTNNRTTLSPKSEKEPNTIRDFPIPKIRDVREFPVLTISNQKKMGEVEFIKVNKDKHSE
SLLGAKFQLQIEKDFSGYKQFVPEGSDVTTKNDGKIYFKALQDGNYKLYEISSPDGYIEV
KTKPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITNTPKRPPGV

SEQ ID NO. 4115: SAG0649 FROM M781 GBS TYPE III STRAIN
GKVIVKKTGDTATPLGKATFVLKNDNDKSETSHETVEGSGKATFENIKPGDYTLREETAP
IGYKKTDKTWKVKVADNGAXIIEGMDADKAEKRKEVLNAQYPKSAIYEDTKENYPLVNVE
GSKVGEQYKALNPINGKDGRREIAEGWLSKKITGVNDLDKNKYKIELTVEGKTTVETKEL
NQPLDVVVLLDNSNSMNNERANNSQRALKAGEAVEKLIDKITSNKDNRVALVTYASTIFD
GTEATVSKGVADQNGKALNDSVSWDYHKTTFTATTHNYSYLNLTNDANEVNILKSRIPKE
AEHINGDRTLYQFGATFTQKALMKANEILETQSSNARKKLIFHVTDGVPTMSYAINFNPY
ISTSYQNQFNSFLNKIPDRSGILQEDFIINGDDYQIVKGDGESFKLFSDRKVPVTGGTTQ
AAYRVPQNQLSVMSNEGYAINSGYIYLYWRDYNWVYPFDPKTKKVSATKQIKTHGEPTTL
YFNGNIRPKGYDIFTVGIGVNGDPGATPLEAEKFMQSISSKTENYTNVDDTNKIYDELNK
YFKTIVEEKHSIVDGNVTDPMGEMIEFQLKNGQSFTHDDYVLVGNDGSQLKNGVALGGPN
SDGGILKDVTVTYDKTSQTIKINHLNLGSGQKVVLTYDVRLKDNYISNKFYNTNNRTTLS
PKSEKEPNTIRDFPIPKIRDVREFPVLTISNQKKMGEVEFIKVNKDKHSESLLGAKFQLQ
IEKDFSGYKQFVPEGSDVTTKNDGKIYFKALQDGNYKLYEISSPDGYIEVKTKPVVTFTI
QNGEVTNLKADPNANKNQIGYLEGNGKHLITNTPKRPPGV

SEQ ID NO. 4117: SAG0649 FROM CJB110 GBS NONTYPEABLE STRAIN
GETQDTNQALGKVIVKKTGDNATPLGKATFVLKNDNDKSETSHETVEGSGXATFENIKPG
DYTLREETAPIGYKKTDKTWKVKVADNGATIIEGMDADKAEKRKEVLNAQYPKSAIYEDT
KENYPLVNVEGSKVGEQYKALNPINGKDGRREIAEGWLSKKITGVNDLDKNKYKIELTVE
GKTTVETKELNQPLDVVVLLDNSNSMNNERANNSQRALKAGEAVEKLIDKITSNKDNRVA
LVTYASTIFDGTEATVSKGVADQNGKALNDSVSWDYHKTTFTATTHNYSYLNLTNDANEV
NILKSRIPKEAEHINGDRTLYQFGATFTQKALMKANEILETQSSNARKKLIFHVTDGVPT
MSYAINFNPYISTSYQNQFNSFLNKIPDRSGILQEDFIINGDDYQIVKGDGESFKLFSDR
KVPVTGGTTQAAYRVPQNQLSVMSNEGYAINSGYIYLYWRDYNWVYPFDPKTKKVSATKQ
IKTHGEPTTLYFNGNIRPKGYDIFTVGIGVNGDPGATPLEAEKFMQSISSKTENYTNVDD
TNKIYDELNKYFKTIVEEKHSIVDGNVTDPMGEMIEFQLKNGQSFTHDDYVLVGNDGSQL
KNGVALGGPNSDGGILKDVTVTYDKTSQTIKINHLNLGSGQKVVLTYDVRLKDNYISNKF
YNTNNRTTLSPKSEKEPNTIRDFPIPKIRDVREFPVLTISNQKKMGEVEFIKVNKDKHSE
SLLGAKFQLQIEKDFSGYKQFVPEGSDVTTKNDGKIYFKALQDGNYKLYEISSPDGYIEV
KTKPVVTFTIQNGEVTNLKADPNANKNQIGYLEGNGKHLITNTPKRPPGV

SEQ ID NO. 4118: SAG0649 FROM JM9130013 GBS TYPE VIII STRAIN
GETQDTNQALGKVIVKKTGDNATPLGKATFVLKNDNDKSETSHETVEGSGEATFENIKPG
DYTLREETAPIGYKKTDKTWKVKVADNGATIIEGMDADKAEKRKEVLNAQYPKSAIYEDT
KENYPLVNVEGSKVGEQYKALNPINGKDGRREIAEGWLSKKITGVNDLDKNKYKIELTVE
GKTTVETKELNQPLDVVVLLDNSNSMNNERANNSQRALKAGEAVEKLIDKITSNKDNRVA
LVTYASTIFDGTEATVSKGVADQNGKALNDSVSWDYHKTTFTATTHNYSYLNLTNDANEV
NILKSRIPKEAEHINGDRTLYQFGATFTQKALMKANEILETQSSNARKKLIFHVTDGVPT
MSYAINFNPYISTSYQNQFNSFLNKIPDRSGILQEDFIINGDDYQIVKGDGESFKLFSDR
KVPVTGGTTQAAYRVPQNQLSVMSNEGYAINSGYIYLYWRDYNWVYPFDPKTKKVSATKQ
IKTHGEPTTLYFNGNIRPKGYDIFTVGIGVNGDPGATPLEAEKFMQSISSKTENYTNVDD
TNKIYDELNKYFKTIVEEKHSIVDGNVTDPMGEMIEFQLKNGQSFTHDDYVLVGNDGSQL
KNGVALGGPNSDGGILKDVTVTYDKTSQTIKINHLNLGSGQKVVLTYDVRLKDNYISNKF
YNTNNRTTLSPKSEKEPNTIRDFPIPKIRDVREFPVLTISNQKKMGEVEFIKVNKDKHSE
SLLGAKFQLQIKKDFSGYKQFVPEGSDVTTKNDGKIYFKALQDGNYKLYEISSPDGYIEV
KTKPVVTFTIQNGEVTNLKADPNANKNQIGYLE

Table 41: Comparative Sequences relating to SAG0649

```
PRETTY of: /biotmp/msa178371.2{*}   May 12, 2003 09:25  ..

                               1                                                  50
     msa178371.2{104_CJB110}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~ge tqdtnqalGK VIVKKTGDnA
       msa178371.2{104_M781}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~GK VIVKKTGDtA
       msa178371.2{104_COH1}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~ge tqdtnqalGK VIVKKTGDnA
       msa178371.2{104_M732}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~ge tqdtnqalGK VIVKKTGDnA
        msa178371.2{104_090}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~ge tqdtnqalGK VIVKKTGDnA
     msa178371.2{104_18RS21}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~ge tqdtnqalGK VIVKKTGDnA
       msa178371.2{104_2603}   mkkrqkiwrg lsvtllilsq ipfgilvqge tqdtnqalGK VIVKKTGDnA
       msa178371.2{104_A909}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~ge tqdtnqalGK VIVKKTGDnA
  msa178371.2{104_JM9130013}   ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~ge tqdtnqalGK VIVKKTGDnA
                   Consensus   ********** ********** ********-- --------** ********_*

                               51                                                100
     msa178371.2{104_CJB110}   TPLGKATFVL KNDNDKSETS HETVEGSGxA TFENIKPGDY TLREETAPIG
       msa178371.2{104_M781}   TPLGKATFVL KNDNDKSETS HETVEGSGkA TFENIKPGDY TLREETAPIG
       msa178371.2{104_COH1}   TPLGKATFVL KNDNDKSETS HETVEGSGxA TFENIKPGDY TLREETAPIG
       msa178371.2{104_M732}   TPLGKATFVL KNDNDKSETS HETVEGSGeA TFENIKPGDY TLREETAPIG
        msa178371.2{104_090}   TPLGKATFVL KNDNDKSETS HETVEGSGeA TFENIKPGDY TLREETAPIG
     msa178371.2{104_18RS21}   TPLGKATFVL KNDNDKSETS HETVEGSGeA TFENIKPGDY TLREETAPIG
       msa178371.2{104_2603}   TPLGKATFVL KNDNDKSETS HETVEGSGeA TFENIKPGDY TLREETAPIG
       msa178371.2{104_A909}   TPLGKATFVL KNDNDKSETS HETVEGSGeA TFENIKPGDY TLREETAPIG
  msa178371.2{104_JM9130013}   TPLGKATFVL KNDNDKSETS HETVEGSGeA TFENIKPGDY TLREETAPIG
                   Consensus   ********** ********** ********_* ********** **********

                               101                                               150
     msa178371.2{104_CJB110}   YKKTDKTWKV KVADNGAtII EGMDADKAEK RKEVLNAQYP KSAIYEDTKE
       msa178371.2{104_M781}   YKKTDKTWKV KVADNGAxII EGMDADKAEK RKEVLNAQYP KSAIYEDTKE
       msa178371.2{104_COH1}   YKKTDKTWKV KVADNGAtII EGMDADKAEK RKEVLNAQYP KSAIYEDTKE
       msa178371.2{104_M732}   YKKTDKTWKV KVADNGAtII EGMDADKAEK RKEVLNAQYP KSAIYEDTKE
        msa178371.2{104_090}   YKKTDKTWKV KVADNGAtII EGMDADKAEK RKEVLNAQYP KSAIYEDTKE
     msa178371.2{104_18RS21}   YKKTDKTWKV KVADNGAtII EGMDADKAEK RKEVLNAQYP KSAIYEDTKE
       msa178371.2{104_2603}   YKKTDKTWKV KVADNGAtII EGMDADKAEK RKEVLNAQYP KSAIYEDTKE
       msa178371.2{104_A909}   YKKTDKTWKV KVADNGAtII EGMDADKAEK RKEVLNAQYP KSAIYEDTKE
  msa178371.2{104_JM9130013}   YKKTDKTWKV KVADNGAtII EGMDADKAEK RKEVLNAQYP KSAIYEDTKE
                   Consensus   ********** *******_** ********** ********** **********

                               151                                               200
     msa178371.2{104_CJB110}   NYPLVNVEGS KVGEQYKALN PINGKDGRRE IAEGWLSKKi TGVNDLDKNK
       msa178371.2{104_M781}   NYPLVNVEGS KVGEQYKALN PINGKDGRRE IAEGWLSKKi TGVNDLDKNK
       msa178371.2{104_COH1}   NYPLVNVEGS KVGEQYKALN PINGKDGRRE IAEGWLSKKn TGVNDLDKNK
       msa178371.2{104_M732}   NYPLVNVEGS KVGEQYKALN PINGKDGRRE IAEGWLSKKn TGVNDLDKNK
        msa178371.2{104_090}   NYPLVNVEGS KVGEQYKALN PINGKDGRRE IAEGWLSKKi TGVNDLDKNK
     msa178371.2{104_18RS21}   NYPLVNVEGS KVGEQYKALN PINGKDGRRE IAEGWLSKKi TGVNDLDKNK
       msa178371.2{104_2603}   NYPLVNVEGS KVGEQYKALN PINGKDGRRE IAEGWLSKKi TGVNDLDKNK
       msa178371.2{104_A909}   NYPLVNVEGS KVGEQYKALN PINGKDGRRE IAEGWLSKKi TGVNDLDKNK
  msa178371.2{104_JM9130013}   NYPLVNVEGS KVGEQYKALN PINGKDGRRE IAEGWLSKKi TGVNDLDKNK
                   Consensus   ********** ********** ********** *********_ **********

                               201                                               250
     msa178371.2{104_CJB110}   YKIELTVEGK TTVETKELNQ PLDVVVLLDN SNSMNNERAN NSQRALKAGE
       msa178371.2{104_M781}   YKIELTVEGK TTVETKELNQ PLDVVVLLDN SNSMNNERAN NSQRALKAGE
       msa178371.2{104_COH1}   YKIELTVEGK TTVETKELNQ PLDVVVLLDN SNSMNNERAN NSQRALKAGE
       msa178371.2{104_M732}   YKIELTVEGK TTVETKELNQ PLDVVVLLDN SNSMNNERAN NSQRALKAGE
        msa178371.2{104_090}   YKIELTVEGK TTVETKELNQ PLDVVVLLDN SNSMNNERAN NSQRALKAGE
     msa178371.2{104_18RS21}   YKIELTVEGK TTVETKELNQ PLDVVVLLDN SNSMNNERAN NSQRALKAGE
       msa178371.2{104_2603}   YKIELTVEGK TTVETKELNQ PLDVVVLLDN SNSMNNERAN NSQRALKAGE
       msa178371.2{104_A909}   YKIELTVEGK TTVETKELNQ PLDVVVLLDN SNSMNNERAN NSQRALKAGE
  msa178371.2{104_JM9130013}   YKIELTVEGK TTVETKELNQ PLDVVVLLDN SNSMNNERAN NSQRALKAGE
                   Consensus   ********** ********** ********** ********** **********

                               251                                               300
     msa178371.2{104_CJB110}   AVEKLIDKIT SNKDNRVALV TYASTIFDGT EATVSKGVAD QNGKALNDSV
       msa178371.2{104_M781}   AVEKLIDKIT SNKDNRVALV TYASTIFDGT EATVSKGVAD QNGKALNDSV
       msa178371.2{104_COH1}   AVEKLIDKIT SNKDNRVALV TYASTIFDGT EATVSKGVAD QNGKALNDSV
       msa178371.2{104_M732}   AVEKLIDKIT SNKDNRVALV TYASTIFDGT EATVSKGVAD QNGKALNDSV
        msa178371.2{104_090}   AVEKLIDKIT SNKDNRVALV TYASTIFDGT EATVSKGVAD QNGKALNDSV
     msa178371.2{104_18RS21}   AVEKLIDKIT SNKDNRVALV TYASTIFDGT EATVSKGVAD QNGKALNDSV
       msa178371.2{104_2603}   AVEKLIDKIT SNKDNRVALV TYASTIFDGT EATVSKGVAD QNGKALNDSV
       msa178371.2{104_A909}   AVEKLIDKIT SNKDNRVALV TYASTIFDGT EATVSKGVAD QNGKALNDSV
  msa178371.2{104_JM9130013}   AVEKLIDKIT SNKDNRVALV TYASTIFDGT EATVSKGVAD QNGKALNDSV
                   Consensus   ********** ********** ********** ********** **********

                               301                                               350
     msa178371.2{104_CJB110}   SWDYHKTTFT ATTHNYSYLN LTNDANEVNI LKSRIPKEAE HINGDRTLYQ
       msa178371.2{104_M781}   SWDYHKTTFT ATTHNYSYLN LTNDANEVNI LKSRIPKEAE HINGDRTLYQ
       msa178371.2{104_COH1}   SWDYHKTTFT ATTHNYSYLN LTNDANEVNI LKSRIPKEAE HINGDRTLYQ
       msa178371.2{104_M732}   SWDYHKTTFT ATTHNYSYLN LTNDANEVNI LKSRIPKEAE HINGDRTLYQ
        msa178371.2{104_090}   SWDYHKTTFT ATTHNYSYLN LTNDANEVNI LKSRIPKEAE HINGDRTLYQ
     msa178371.2{104_18RS21}   SWDYHKTTFT ATTHNYSYLN LTNDANEVNI LKSRIPKEAE HINGDRTLYQ
       msa178371.2{104_2603}   SWDYHKTTFT ATTHNYSYLN LTNDANEVNI LKSRIPKEAE HINGDRTLYQ
       msa178371.2{104_A909}   SWDYHKTTFT ATTHNYSYLN LTNDANEVNI LKSRIPKEAE HINGDRTLYQ
```

Table 41: Comparative Sequences relating to SAG0649

```
msa178371.2{104_JM9130013}  SWDYHKTTFT ATTHNYSYLN LTNDANEVNI LKSRIPKEAE HINGDRTLYQ
                 Consensus  ********** ********** ********** ********** **********

                            351                                               400
    msa178371.2{104_CJB110}  FGATFTQKAL MKANEILETQ SSNARKKLIF HVTDGVPTMS YAINFNPYIS
      msa178371.2{104_M781}  FGATFTQKAL MKANEILETQ SSNARKKLIF HVTDGVPTMS YAINFNPYIS
      msa178371.2{104_COH1}  FGATFTQKAL MKANEILETQ SSNARKKLIF HVTDGVPTMS YAINFNPYIS
      msa178371.2{104_M732}  FGATFTQKAL MKANEILETQ SSNARKKLIF HVTDGVPTMS YAINFNPYIS
       msa178371.2{104_090}  FGATFTQKAL MKANEILETQ SSNARKKLIF HVTDGVPTMS YAINFNPYIS
    msa178371.2{104_18RS21}  FGATFTQKAL MKANEILETQ SSNARKKLIF HVTDGVPTMS YAINFNPYIS
      msa178371.2{104_2603}  FGATFTQKAL MKANEILETQ SSNARKKLIF HVTDGVPTMS YAINFNPYIS
      msa178371.2{104_A909}  FGATFTQKAL MKANEILETQ SSNARKKLIF HVTDGVPTMS YAINFNPYIS
 msa178371.2{104_JM9130013}  FGATFTQKAL MKANEILETQ SSNARKKLIF HVTDGVPTMS YAINFNPYIS
                 Consensus  ********** ********** ********** ********** **********

                            401                                               450
    msa178371.2{104_CJB110}  TSYQNQFNSF LNKIPDRSGI LQEDFIINGD DYQIVKGDGE SFKLFSDRKV
      msa178371.2{104_M781}  TSYQNQFNSF LNKIPDRSGI LQEDFIINGD DYQIVKGDGE SFKLFSDRKV
      msa178371.2{104_COH1}  TSYQNQFNSF LNKIPDRSGI LQEDFIINGD DYQIVKGDGE SFKLFSDRKV
      msa178371.2{104_M732}  TSYQNQFNSF LNKIPDRSGI LQEDFIINGD DYQIVKGDGE SFKLFSDRKV
       msa178371.2{104_090}  TSYQNQFNSF LNKIPDRSGI LQEDFIINGD DYQIVKGDGE SFKLFSDRKV
    msa178371.2{104_18RS21}  TSYQNQFNSF LNKIPDRSGI LQEDFIINGD DYQIVKGDGE SFKLFSDRKV
      msa178371.2{104_2603}  TSYQNQFNSF LNKIPDRSGI LQEDFIINGD DYQIVKGDGE SFKLFSDRKV
      msa178371.2{104_A909}  TSYQNQFNSF LNKIPDRSGI LQEDFIINGD DYQIVKGDGE SFKLFSDRKV
 msa178371.2{104_JM9130013}  TSYQNQFNSF LNKIPDRSGI LQEDFIINGD DYQIVKGDGE SFKLFSDRKV
                 Consensus  ********** ********** ********** ********** **********

                            451                                               500
    msa178371.2{104_CJB110}  PVTGGTTQAA YRVPQNQLSV MSNEGYAINS GYIYLYWRDY NWVYPFDPKT
      msa178371.2{104_M781}  PVTGGTTQAA YRVPQNQLSV MSNEGYAINS GYIYLYWRDY NWVYPFDPKT
      msa178371.2{104_COH1}  PVTGGTTQAA YRVPQNQLSV MSNEGYAINS GYIYLYWRDY NWVYPFDPKT
      msa178371.2{104_M732}  PVTGGTTQAA YRVPQNQLSV MSNEGYAINS GYIYLYWRDY NWVYPFDPKT
       msa178371.2{104_090}  PVTGGTTQAA YRVPQNQLSV MSNEGYAINS GYIYLYWRDY NWVYPFDPKT
    msa178371.2{104_18RS21}  PVTGGTTQAA YRVPQNQLSV MSNEGYAINS GYIYLYWRDY NWVYPFDPKT
      msa178371.2{104_2603}  PVTGGTTQAA YRVPQNQLSV MSNEGYAINS GYIYLYWRDY NWVYPFDPKT
      msa178371.2{104_A909}  PVTGGTTQAA YRVPQNQLSV MSNEGYAINS GYIYLYWRDY NWVYPFDPKT
 msa178371.2{104_JM9130013}  PVTGGTTQAA YRVPQNQLSV MSNEGYAINS GYIYLYWRDY NWVYPFDPKT
                 Consensus  ********** ********** ********** ********** **********

                            501                                               550
    msa178371.2{104_CJB110}  KKVSATKQIK THGEPTTLYF NGNIRPKGYD IFTVGIGVNG DPGATPLEAE
      msa178371.2{104_M781}  KKVSATKQIK THGEPTTLYF NGNIRPKGYD IFTVGIGVNG DPGATPLEAE
      msa178371.2{104_COH1}  KKVSATKQIK THGEPTTLYF NGNIRPKGYD IFTVGIGVNG DPGATPLEAE
      msa178371.2{104_M732}  KKVSATKQIK THGEPTTLYF NGNIRPKGYD IFTVGIGVNG DPGATPLEAE
       msa178371.2{104_090}  KKVSATKQIK THGEPTTLYF NGNIRPKGYD IFTVGIGVNG DPGATPLEAE
    msa178371.2{104_18RS21}  KKVSATKQIK THGEPTTLYF NGNIRPKGYD IFTVGIGVNG DPGATPLEAE
      msa178371.2{104_2603}  KKVSATKQIK THGEPTTLYF NGNIRPKGYD IFTVGIGVNG DPGATPLEAE
      msa178371.2{104_A909}  KKVSATKQIK THGEPTTLYF NGNIRPKGYD IFTVGIGVNG DPGATPLEAE
 msa178371.2{104_JM9130013}  KKVSATKQIK THGEPTTLYF NGNIRPKGYD IFTVGIGVNG DPGATPLEAE
                 Consensus  ********** ********** ********** ********** **********

                            551                                               600
    msa178371.2{104_CJB110}  KFMQSISSKT ENYTNVDDTN KIYDELNKYF KTIVEEKHSI VDGNVTDPMG
      msa178371.2{104_M781}  KFMQSISSKT ENYTNVDDTN KIYDELNKYF KTIVEEKHSI VDGNVTDPMG
      msa178371.2{104_COH1}  KFMQSISSKT ENYTNVDDTN KIYDELNKYF KTIVEEKHSI VDGNVTDPMG
      msa178371.2{104_M732}  KFMQSISSKT ENYTNVDDTN KIYDELNKYF KTIVEEKHSI VDGNVTDPMG
       msa178371.2{104_090}  KFMQSISSKT ENYTNVDDTN KIYDELNKYF KTIVEEKHSI VDGNVTDPMG
    msa178371.2{104_18RS21}  KFMQSISSKT ENYTNVDDTN KIYDELNKYF KTIVEEKHSI VDGNVTDPMG
      msa178371.2{104_2603}  KFMQSISSKT ENYTNVDDTN KIYDELNKYF KTIVEEKHSI VDGNVTDPMG
      msa178371.2{104_A909}  KFMQSISSKT ENYTNVDDTN KIYDELNKYF KTIVEEKHSI VDGNVTDPMG
 msa178371.2{104_JM9130013}  KFMQSISSKT ENYTNVDDTN KIYDELNKYF KTIVEEKHSI VDGNVTDPMG
                 Consensus  ********** ********** ********** ********** **********

                            601                                               650
    msa178371.2{104_CJB110}  EMIEFQLKNG QSFTHDDYVL VGNDGSQLKN GVALGGPNSD GGILKDVTVT
      msa178371.2{104_M781}  EMIEFQLKNG QSFTHDDYVL VGNDGSQLKN GVALGGPNSD GGILKDVTVT
      msa178371.2{104_COH1}  EMIEFQLKNG QSFTHDDYVL VGNDGSQLKN GVALGGPNSD GGILKDVTVT
      msa178371.2{104_M732}  EMIEFQLKNG QSFTHDDYVL VGNDGSQLKN GVALGGPNSD GGILKDVTVT
       msa178371.2{104_090}  EMIEFQLKNG QSFTHDDYVL VGNDGSQLKN GVALGGPNSD GGILKDVTVT
    msa178371.2{104_18RS21}  EMIEFQLKNG QSFTHDDYVL VGNDGSQLKN GVALGGPNSD GGILKDVTVT
      msa178371.2{104_2603}  EMIEFQLKNG QSFTHDDYVL VGNDGSQLKN GVALGGPNSD GGILKDVTVT
      msa178371.2{104_A909}  EMIEFQLKNG QSFTHDDYVL VGNDGSQLKN GVALGGPNSD GGILKDVTVT
 msa178371.2{104_JM9130013}  EMIEFQLKNG QSFTHDDYVL VGNDGSQLKN GVALGGPNSD GGILKDVTVT
                 Consensus  ********** ********** ********** ********** **********

                            651                                               700
    msa178371.2{104_CJB110}  YDKTSQTIKI NHLNLGSGQK VVLTYDVRLK DNYISNKFYN TNNRTTLSPK
      msa178371.2{104_M781}  YDKTSQTIKI NHLNLGSGQK VVLTYDVRLK DNYISNKFYN TNNRTTLSPK
      msa178371.2{104_COH1}  YDKTSQTIKI NHLNLGSGQK VVLTYDVRLK DNYISNKFYN TNNRTTLSPK
      msa178371.2{104_M732}  YDKTSQTIKI NHLNLGSGQK VVLTYDVRLK DNYISNKFYN TNNRTTLSPK
       msa178371.2{104_090}  YDKTSQTIKI NHLNLGSGQK VVLTYDVRLK DNYISNKFYN TNNRTTLSPK
    msa178371.2{104_18RS21}  YDKTSQTIKI NHLNLGSGQK VVLTYDVRLK DNYISNKFYN TNNRTTLSPK
      msa178371.2{104_2603}  YDKTSQTIKI NHLNLGSGQK VVLTYDVRLK DNYISNKFYN TNNRTTLSPK
```

Table 41: Comparative Sequences relating to SAG0649

```
     msa178371.2{104_A909}  YDKTSQTIKI NHLNLGSGQK VVLTYDVRLK DNYISNKFYN TNNRTTLSPK
msa178371.2{104_JM9130013}  YDKTSQTIKI NHLNLGSGQK VVLTYDVRLK DNYISNKFYN TNNRTTLSPK
                 Consensus  ********** ********** ********** ********** **********

                            701                                              750
   msa178371.2{104_CJB110}  SEKEPNTIRD FPIPKIRDVR EFPVLTISNQ KKMGEVEFIK VNKDKHSESL
     msa178371.2{104_M781}  SEKEPNTIRD FPIPKIRDVR EFPVLTISNQ KKMGEVEFIK VNKDKHSESL
     msa178371.2{104_COH1}  SEKEPNTIRD FPIPKIRDVR EFPVLTISNQ KKMGEVEFIK VNKDKHSESL
     msa178371.2{104_M732}  SEKEPNTIRD FPIPKIRDVR EFPVLTISNQ KKMGEVEFIK VNKDKHSESL
      msa178371.2{104_090}  SEKEPNTIRD FPIPKIRDVR EFPVLTISNQ KKMGEVEFIK VNKDKHSESL
   msa178371.2{104_18RS21}  SEKEPNTIRD FPIPKIRDVR EFPVLTISNQ KKMGEVEFIK VNKDKHSESL
     msa178371.2{104_2603}  SEKEPNTIRD FPIPKIRDVR EFPVLTISNQ KKMGEVEFIK VNKDKHSESL
     msa178371.2{104_A909}  SEKEPNTIRD FPIPKIRDVR EFPVLTISNQ KKMGEVEFIK VNKDKHSESL
msa178371.2{104_JM9130013}  SEKEPNTIRD FPIPKIRDVR EFPVLTISNQ KKMGEVEFIK VNKDKHSESL
                 Consensus  ********** ********** ********** ********** **********

                            751                                              800
   msa178371.2{104_CJB110}  LGAKFQLQIe KDFSGYKQFV PEGSDVTTKN DGKIYFKALQ DGNYKLYEIS
     msa178371.2{104_M781}  LGAKFQLQIe KDFSGYKQFV PEGSDVTTKN DGKIYFKALQ DGNYKLYEIS
     msa178371.2{104_COH1}  LGAKFQLQIe KDFSGYKQFV PEGSDVTTKN DGKIYFKALQ DGNYKLYEIS
     msa178371.2{104_M732}  LGAKFQLQIe KDFSGYKQFV PEGSDVTTKN DGKIYFKALQ DGNYKLYEIS
      msa178371.2{104_090}  LGAKFQLQIe KDFSGYKQFV PEGSDVTTKN DGKIYFKALQ DGNYKLYEIS
   msa178371.2{104_18RS21}  LGAKFQLQIe KDFSGYKQFV PEGSDVTTKN DGKIYFKALQ DGNYKLYEIS
     msa178371.2{104_2603}  LGAKFQLQIe KDFSGYKQFV PEGSDVTTKN DGKIYFKALQ DGNYKLYEIS
     msa178371.2{104_A909}  LGAKFQLQIe KDFSGYKQFV PEGSDVTTKN DGKIYFKALQ DGNYKLYEIS
msa178371.2{104_JM9130013}  LGAKFQLQIk KDFSGYKQFV PEGSDVTTKN DGKIYFKALQ DGNYKLYEIS
                 Consensus  *********- ********** ********** ********** **********

                            801                                              850
   msa178371.2{104_CJB110}  SPDGYIEVKT KPVVTFTIQN GEVTNLKADP NANKNQIGYL Egngkhlitn
     msa178371.2{104_M781}  SPDGYIEVKT KPVVTFTIQN GEVTNLKADP NANKNQIGYL Egngkhlitn
     msa178371.2{104_COH1}  SPDGYIEVKT KPVVTFTIQN GEVTNLKADP NANKNQIGYL Egngkhlitn
     msa178371.2{104_M732}  SPDGYIEVKT KPVVTFTIQN GEVTNLKADP NANKNQIGYL Egngkhlitn
      msa178371.2{104_090}  SPDGYIEVKT KPVVTFTIQN GEVTNLKADP NANKNQIGYL Egngkhlitn
   msa178371.2{104_18RS21}  SPDGYIEVKT KPVVTFTIQN GEVTNLKADP NANKNQIGYL Egngkhlitn
     msa178371.2{104_2603}  SPDGYIEVKT KPVVTFTIQN GEVTNLKADP NANKNQIGYL Egngkhlitn
     msa178371.2{104_A909}  SPDGYIEVKT KPVVTFTIQN GEVTNLKADP NANKNQIGYL Egngkhlitn
msa178371.2{104_JM9130013}  SPDGYIEVKT KPVVTFTIQN GEVTNLKADP NANKNQIGYL E~~~~~~~~~
                 Consensus  ********** ********** ********** ********** *---------

                            851                                   890
   msa178371.2{104_CJB110}  tpkrppgv~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     msa178371.2{104_M781}  tpkrppgv~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     msa178371.2{104_COH1}  tpkrppgv~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     msa178371.2{104_M732}  tpkrppgv~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      msa178371.2{104_090}  tpkrppgv~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   msa178371.2{104_18RS21}  tpkrppgv~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     msa178371.2{104_2603}  tpkrppgvfp ktggigtivy ilvgstfmil ticsfrrkql
     msa178371.2{104_A909}  tpkrppgv~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
msa178371.2{104_JM9130013}  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
                 Consensus  --------** ********** ********** **********
```